



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number: 166704

TO: Nita M Minnifield
Location: REM/3C01/3C18
Art Unit: 1645
Thursday, September 29, 2005
Case Serial Number: 09/298523

From: Barb O'Bryen
Location: Biotech-Chem Library
Remsen 1a69
Phone: 571-272-2518 *BOB*

barbara.obryen@uspto.gov

Search Notes

*Reviewed
9/05*

THIS PAGE IS BLANK

STIC-Biotech/ChemLib

166904

From: Minnifield, Nita
Sent: Friday, September 23, 2005 12:37 PM
To: STIC-Biotech/ChemLib
Subject: sequence search

09/298523

STIC

Please do a commercial and interference sequence search on SEQ
ID NO: 13 of this application.

Please provide a paper copy.

Thanks,
Minnifield
71976
Art Unit 1645
Office REM-3C01
Mailbox REM-3C18
571-272-0860

RECEIVED
SEP 23 2005
TECH/CHEM. DIVISION
(STIC)

Searcher: _____
Searcher Phone: _____
Date Searcher Picked up: _____
Date completed: _____
Searcher Prep Time: _____
Online Time: _____

Type of Search
NA# _____ AA# _____
S/L: _____ Oligomer: _____
Encode/Transl: _____
Structure #: _____ Text: _____
Inventor: _____ Litigation: _____

Vendors and cost where applicable
STN: _____
DIALOG: _____
QUESTEL/ORBIT: _____
LEXIS/NEXIS: _____
SEQUENCE SYSTEM: _____
WWW/Internet: _____
Other (Specify): _____

THIS PAGE IS BLANK

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 28, 2005, 18:42:50 ; Search time 173 Seconds
(without alignments)
1206.960 Million cell updates/sec

Title: US-09-298-523D-13

Perfect score: 2546

Sequence: 1 MFASKSRKVHYSIRKFSVG.....NRLTQQQPKTEKPAQSTP 513

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1826554 seqs, 407025358 residues

Total number of hits satisfying chosen parameters: 1826554

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA.*

1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US10E_PUBCOMB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
19: /cgn2_6/ptodata/2/pubpaa/US11A_PUBCOMB.pep.*
20: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
21: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
22: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	2546	100.0	693	10	US-09-769-787-185
2	2546	100.0	693	15	US-10-333-120A-10
3	2546	100.0	693	17	US-10-472-928-4598
4	2516	98.8	511	9	US-09-748-875-13
5	2516	98.8	511	10	US-09-298-523B-13
6	2516	98.8	511	18	US-10-341-201-13
7	2505	98.4	513	9	US-09-748-875-12
8	2505	98.4	513	10	US-09-298-523B-12
9	2505	98.4	513	18	US-10-341-201-12
10	2366	92.9	655	10	US-09-056-019-2
11			655	16	US-10-751-702-2

12	2268	89.1	460	10	US-09-056-019-38
13	2268	89.1	460	16	US-10-751-702-38
14	2251	88.4	459	10	US-09-056-019-39
15	2251	88.4	459	16	US-10-751-702-39
16	2210	86.8	446	14	US-10-254-995-9
17	2210	86.8	446	20	US-11-062-080-9
18	2193	86.1	446	14	US-10-254-995-6
19	2193	86.1	446	20	US-11-062-080-6
20	2119	83.2	451	9	US-09-748-875-67
21	2119	83.2	451	10	US-09-298-523B-67
22	2119	83.2	451	18	US-10-341-201-67
23	2117	83.2	428	10	US-09-056-019-24
24	2117	83.2	428	16	US-10-751-702-24
25	1995	78.4	406	10	US-09-056-019-1
26	1995	78.4	406	16	US-10-751-702-1
27	1796.5	70.6	487	15	US-10-333-120A-9
28	1787.5	70.2	487	9	US-09-748-875-9
29	1787.5	70.2	487	10	US-09-298-523B-9
30	1787.5	70.2	487	18	US-10-341-201-9
31	1771.5	69.6	487	9	US-09-748-875-66
32	1771.5	69.6	487	10	US-09-298-523B-66
33	1771.5	69.6	487	18	US-10-341-201-66
34	1756	69.0	701	15	US-10-282-122A-74228
35	1752	68.8	701	15	US-10-333-120A-7
36	1744	68.5	581	9	US-09-748-875-56
37	1744	68.5	581	10	US-09-298-523B-56
38	1744	68.5	581	18	US-10-341-201-56
39	1741	68.4	483	9	US-09-748-875-10
40	1741	68.4	483	10	US-09-298-523B-10
41	1741	68.4	483	18	US-10-341-201-10
42	1726	67.8	481	9	US-09-748-875-6
43	1726	67.8	481	10	US-09-298-523B-6
44	1726	67.8	481	18	US-10-341-201-6
45	1705.5	67.0	496	9	US-09-748-875-4

ALIGNMENTS

RESULT 1

US-09-769-787-185
; Sequence 185, Application US/09769787
; Publication No. US20030091577A1
; GENERAL INFORMATION:
; APPLICANT: Microbial Technics Limited
; APPLICANT: Gilbert, Christophe FG
; APPLICANT: Hansbro, Philip M
; TITLE OF INVENTION: Proteins
; FILE REFERENCE: PWC/P21129WO
; CURRENT APPLICATION NUMBER: US/09/769,787
; CURRENT FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: GB 9816337.1
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: US 60/125164
; PRIOR FILING DATE: 1999-03-19
; NUMBER OF SEQ ID NOS: 388.
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 185
; LENGTH: 693
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-769-787-185

Query Match	100.0%;	Score 2546;	DB 10;	Length 693;
Best Local Similarity	100.0%;	Pred. No. 1.7e-107;		
Matches 513;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	MFASKSRKVHYSIRKFSVGVS	VVSLVMSGVVHATENEGATQVPTSSNRANSQA	BQ 60
Db	1	MFASKSRKVHYSIRKFSVGVS	VVSLVMSGVVHATENEGATQVPTSSNRANSQA	BQ 60
Qy	61	GEOPKLLSEDKARKEVEYKIKVGSYAKSTKKRHTITVALNNELNINIKIV	120	

Db 61 GEQPKLDSERDKARKEVEEYVKKIVGESYAKSTKGRHTITVALVNLNNIKNEYLNKIV 120
QY 121 ESTSESQILIMBSRSKVDKAVSKFEKDSSSSSSSSTKPEASDTAKPNKPTPEGEKV 180
Db 121 ESTSESQILIMBSRSKVDKAVSKFEKDSSSSSSSSTKPEASDTAKPNKPTPEGEKV 180
QY 181 AEAKKKVEAEAKKADQKEEDRRNYPTITYKTLELEIAESDVEYKKAELVLVKYKANEPR 240
Db 181 AEAKKKVEAEAKKADQKEEDRRNYPTITYKTLELEIAESDVEYKKAELVLVKYKANEPR 240
QY 241 DEQIKIQAIAEVESKQAEATRLKKIKTDREAEAEAKRRADAKAQGPKGKRAKGVPGEL 300
Db 241 DEQIKIQAIAEVESKQAEATRLKKIKTDREAEAEAKRRADAKAQGPKGKRAKGVPGEL 300
QY 301 ATPDKKENDAKSSDSSVGEETLPSLKPCKKVAEAKKVEAEAKKADQKEEDRRNYPT 360
Db 301 ATPDKKENDAKSSDSSVGEETLPSLKPCKKVAEAKKVEAEAKKADQKEEDRRNYPT 360
QY 361 NTYKTLELEIAESDVEYKKAELVLKKEAEKPRNEEKVKQAIAEVESKKAETRLKIKT 420
Db 361 NTYKTLELEIAESDVEYKKAELVLKKEAEKPRNEEKVKQAIAEVESKKAETRLKIKT 420
QY 421 DRKKAEEAKKAAEEDKVKKEPAEQPAPAPAKAEKPAAPAKPENPAEQPKAEKPADQ 480
Db 421 DRKKAEEAKKAAEEDKVKKEPAEQPAPAPAKAEKPAAPAKPENPAEQPKAEKPADQ 480
QY 481 AEEDYARRSEEEYNRLTQQQPPKTEKPAQPSTP 513
Db 481 AEEDYARRSEEEYNRLTQQQPPKTEKPAQPSTP 513

RESULT 2

US-10-333-120A-10
; Sequence 10, Application US/10333120A
; Publication No. US20040091495A1
; GENERAL INFORMATION:
; APPLICANT: Janulczyk, Robert
; APPLICANT: Iannelli, Francesco
; APPLICANT: Sjoelhm, Anders
; APPLICANT: Pozzi, Gianni
; APPLICANT: Bjorck, Lars
; TITLE OF INVENTION: PROTEIN
; CURRENT APPLICATION NUMBER: 13519PCTUS
; CURRENT FILING DATE: 2002-01-14
; PRIOR APPLICATION NUMBER: PCT/E901/08409
; PRIOR FILING DATE: 2001-07-19
; PRIOR FILING DATE: 2000-07-20
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 693
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-333-120A-10

Query Match 100.0%; Score 2546; DB 15; Length 693;
Best Local Similarity 100.0%; Pred. No. 1.7e-107;
Matches 513; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MFASKSERKVHYSIRKFSVGVASVVVASLVMSVVAHATENEGATQVPTSSNRANESQAEQ 60
Db 1 MFASKSERKVHYSIRKFSVGVASVVVASLVMSVVAHATENEGATQVPTSSNRANESQAEQ 60
QY 61 GEQPKLDSERDKARKEVEEYVKKIVGESYAKSTKGRHTITVALVNLNNIKNEYLNKIV 120
Db 61 GEQPKLDSERDKARKEVEEYVKKIVGESYAKSTKGRHTITVALVNLNNIKNEYLNKIV 120
QY 121 ESTSESQILIMBSRSKVDKAVSKFEKDSSSSSSSSTKPEASDTAKPNKPTPEGEKV 180
Db 121 ESTSESQILIMBSRSKVDKAVSKFEKDSSSSSSSSTKPEASDTAKPNKPTPEGEKV 180

QY 181 AEAKKKVEAEAKKADQKEEDRRNYPTITYKTLELEIAESDVEYKKAELVLVKYKANEPR 240
Db 181 AEAKKKVEAEAKKADQKEEDRRNYPTITYKTLELEIAESDVEYKKAELVLVKYKANEPR 240
QY 241 DEQIKIQAIAEVESKQAEATRLKKIKTDREAEAEAKRRADAKAQGPKGKRAKGVPGEL 300
Db 241 DEQIKIQAIAEVESKQAEATRLKKIKTDREAEAEAKRRADAKAQGPKGKRAKGVPGEL 300
QY 301 ATPDKKENDAKSSDSSVGEETLPSLKPCKKVAEAKKVEAEAKKADQKEEDRRNYPT 360
Db 301 ATPDKKENDAKSSDSSVGEETLPSLKPCKKVAEAKKVEAEAKKADQKEEDRRNYPT 360
QY 361 NTYKTLELEIAESDVEYKKAELVLKKEAEKPRNEEKVKQAIAEVESKKAETRLKIKT 420
Db 361 NTYKTLELEIAESDVEYKKAELVLKKEAEKPRNEEKVKQAIAEVESKKAETRLKIKT 420
QY 421 DRKKAEEAKKAAEEDKVKKEPAEQPAPAPAKAEKPAAPAKPENPAEQPKAEKPADQ 480
Db 421 DRKKAEEAKKAAEEDKVKKEPAEQPAPAPAKAEKPAAPAKPENPAEQPKAEKPADQ 480
QY 481 AEEDYARRSEEEYNRLTQQQPPKTEKPAQPSTP 513
Db 481 AEEDYARRSEEEYNRLTQQQPPKTEKPAQPSTP 513

RESULT 3

US-10-472-928-4598
; Sequence 4598, Application US/10472928
; Publication No. US20050020813A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SpA
; TITLE OF INVENTION: STREPTOCOCCUS PNEUMONIAE PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE: P026926WO
; CURRENT APPLICATION NUMBER: US/10/472,928
; CURRENT FILING DATE: 2003-09-26
; PRIOR APPLICATION NUMBER: GB-0107658.7
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 4979
; SOFTWARE: SeqWin99, version 1.03
; SEQ ID NO 4598
; LENGTH: 693
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
; FEATURE:
; OTHER INFORMATION: choline binding protein A (cbpA)
; OTHER INFORMATION: Cellular location: outside
; OTHER INFORMATION: Feature of note: WY motif
; OTHER INFORMATION: Similar to strain R6 sequence 15904036 (0.E+01)
US-10-472-928-4598

Query Match 100.0%; Score 2546; DB 17; Length 693;
Best Local Similarity 100.0%; Pred. No. 1.7e-107;
Matches 513; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MFASKSERKVHYSIRKFSVGVASVVVASLVMSVVAHATENEGATQVPTSSNRANESQAEQ 60
Db 1 MFASKSERKVHYSIRKFSVGVASVVVASLVMSVVAHATENEGATQVPTSSNRANESQAEQ 60
QY 61 GEQPKLDSERDKARKEVEEYVKKIVGESYAKSTKGRHTITVALVNLNNIKNEYLNKIV 120
Db 61 GEQPKLDSERDKARKEVEEYVKKIVGESYAKSTKGRHTITVALVNLNNIKNEYLNKIV 120
QY 121 ESTSESQILIMBSRSKVDKAVSKFEKDSSSSSSSSTKPEASDTAKPNKPTPEGEKV 180
Db 121 ESTSESQILIMBSRSKVDKAVSKFEKDSSSSSSSSTKPEASDTAKPNKPTPEGEKV 180
QY 181 AEAKKKVEAEAKKADQKEEDRRNYPTITYKTLELEIAESDVEYKKAELVLVKYKANEPR 240
Db 181 AEAKKKVEAEAKKADQKEEDRRNYPTITYKTLELEIAESDVEYKKAELVLVKYKANEPR 240
QY 241 DEQIKIQAIAEVESKQAEATRLKKIKTDREAEAEAKRRADAKAQGPKGKRAKGVPGEL 300

```
Db 241 DEQIKQAEAVESKQAEATRLKKIKTDREAEAEAEAKRADAKEQKPKGRAKGVPGEL 300
Qy 301 ATPDKKENDAKSSDSSVGEETLPSLSLPEKKVAEAEKKVEPAKKAEDQKEEDRRNYPT 360
Db 301 ATPDKKENDAKSSDSSVGEETLPSLSLPEKKVAEAEKKVEPAKKAEDQKEEDRRNYPT 360
Qy 361 NTYKTLLEIAESDVVEVKKAELVLKBEAKPRNEEKVKQAKAEVSKKAEATRLLEIKT 420
Db 361 NTYKTLLEIAESDVVEVKKAELVLKBEAKPRNEEKVKQAKAEVSKKAEATRLLEIKT 420
Qy 421 DRKKAEEBAKKAABEDKVKKEPAQPOPPAPAPAKAEKPAPAPKPNPAEQPKAEKPADOQ 480
Db 421 DRKKAEEBAKKAABEDKVKKEPAQPOPPAPAPAKAEKPAPAPKPNPAEQPKAEKPADOQ 480
Qy 481 AEEDYARRSEBEYNRLTQQQPPKTEKPAQSTP 513
Db 481 AEEDYARRSEBEYNRLTQQQPPKTEKPAQSTP 513

RESULT 4
US-09-748-875-13
; Sequence 13, Application US/09748875
; Publication No. US20010016200A1
; GENERAL INFORMATION:
; APPLICANT: BRILES et al.
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEIN C (PSPC), EPITOPIC REGIONS
; FILE REFERENCE: 454312-3140
; CURRENT APPLICATION NUMBER: US/09/748, 875
; PRIOR FILING DATE: 2000-12-26
; PRIOR APPLICATION NUMBER: 09/298, 523
; PRIOR FILING DATE: 1999-04-23
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 511
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-748-875-13

Query Match 98.8%; Score 2516; DB 9; Length 511;
Best Local Similarity 99.6%; Pred. No. 2.9e-106;
Matches 511; Conservative 0; Mismatches 0; Indels 2; Gaps 2;

Qy 1 MFASKSRKVVHYSIRKFSVGVASVVVAVSLVMSVVHATENEGATQVPTSSNRANESQAEQ 60
Db 1 MFASKSRKVVHYSIRKFSVG-ASVVAVSLVMSVVHATENEGATQVPTSSNRANESQAEQ 59
Qy 61 GEQPKLDSERDKARKEVEYVKKIVGESYAKSTKKRHTITVALVNLNLIKNEYLNKIV 120
Db 60 GEQPKLDSERDKARKEVEYVKKIVGESYAKSTKKRHTITVALVNLNLIKNEYLNKIV 119
Qy 121 ESTSESQILMMESRSKVDVAVSKFEKDSSSSSSSSTKPEASDTAKPNKPTPEGBKV 180
Db 120 ESTSESQILMMESRSKVDVAVSKFEKDSSSSSSSSTKPEASDTAKPNKPTPEGBKV 179
Qy 181 AEAKKKVEAEAKKADQKEEDRRNYPTITYKTLLEIAESDVVEVKKAELVLKVKANEP 240
Db 180 AEAKKKVEEA-KKADQKEEDRRNYPTITYKTLLEIAESDVVEVKKAELVLKVKANEP 238
Qy 241 DEQIKQAEAVESKQAEATRLKKIKTDREAEAEAEAKRADAKEQKPKGRAKGVPGEL 300
Db 239 DEQIKQAEAVESKQAEATRLKKIKTDREAEAEAEAKRADAKEQKPKGRAKGVPGEL 298
Qy 301 ATPDKKENDAKSSDSSVGEETLPSLSLPEKKVAEAEKKVEPAKKAEDQKEEDRRNYPT 360
Db 299 ATPDKKENDAKSSDSSVGEETLPSLSLPEKKVAEAEKKVEPAKKAEDQKEEDRRNYPT 358
Qy 361 NTYKTLLEIAESDVVEVKKAELVLKBEAKPRNEEKVKQAKAEVSKKAEATRLLEIKT 420
Db 359 NTYKTLLEIAESDVVEVKKAELVLKBEAKPRNEEKVKQAKAEVSKKAEATRLLEIKT 418
Qy 421 DRKKAEEBAKKAABEDKVKKEPAQPOPPAPAPAKAEKPAPAPKPNPAEQPKAEKPADOQ 480
Db 419 DRKKAEEBAKKAABEDKVKKEPAQPOPPAPAPAKAEKPAPAPKPNPAEQPKAEKPADOQ 478
Qy 481 AEEDYARRSEBEYNRLTQQQPPKTEKPAQSTP 513
Db 479 AEEDYARRSEBEYNRLTQQQPPKTEKPAQSTP 511

RESULT 6
US-10-341-201-13
; Sequence 13, Application US/10341201
; Publication No. US20050196405A1
; GENERAL INFORMATION:
```

```
Db 419 DRKKAEEBAKKAABEDKVKKEPAQPOPPAPAPAKAEKPAPAPKPNPAEQPKAEKPADOQ 478
Qy 481 AEEDYARRSEBEYNRLTQQQPPKTEKPAQSTP 513
Db 479 AEEDYARRSEBEYNRLTQQQPPKTEKPAQSTP 511

RESULT 5
US-09-298-523B-13
; Sequence 13, Application US/09298523B
; Publication No. US20030059438A1
; GENERAL INFORMATION:
; APPLICANT: BRILES et al.
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEIN C (PSPC), EPITOPIC REGIONS
; FILE REFERENCE: 454312-3140
; CURRENT APPLICATION NUMBER: US/09/298, 523B
; CURRENT FILING DATE: 1999-04-23
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 511
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-298-523B-13

Query Match 98.8%; Score 2516; DB 10; Length 511;
Best Local Similarity 99.6%; Pred. No. 2.9e-106;
Matches 511; Conservative 0; Mismatches 0; Indels 2; Gaps 2;

Qy 1 MFASKSRKVVHYSIRKFSVGVASVVVAVSLVMSVVHATENEGATQVPTSSNRANESQAEQ 60
Db 1 MFASKSRKVVHYSIRKFSVG-ASVVAVSLVMSVVHATENEGATQVPTSSNRANESQAEQ 59
Qy 61 GEQPKLDSERDKARKEVEYVKKIVGESYAKSTKKRHTITVALVNLNLIKNEYLNKIV 120
Db 60 GEQPKLDSERDKARKEVEYVKKIVGESYAKSTKKRHTITVALVNLNLIKNEYLNKIV 119
Qy 121 ESTSESQILMMESRSKVDVAVSKFEKDSSSSSSSSTKPEASDTAKPNKPTPEGBKV 180
Db 120 ESTSESQILMMESRSKVDVAVSKFEKDSSSSSSSSTKPEASDTAKPNKPTPEGBKV 179
Qy 181 AEAKKKVEAEAKKADQKEEDRRNYPTITYKTLLEIAESDVVEVKKAELVLKVKANEP 240
Db 180 AEAKKKVEEA-KKADQKEEDRRNYPTITYKTLLEIAESDVVEVKKAELVLKVKANEP 238
Qy 241 DEQIKQAEAVESKQAEATRLKKIKTDREAEAEAEAKRADAKEQKPKGRAKGVPGEL 300
Db 239 DEQIKQAEAVESKQAEATRLKKIKTDREAEAEAEAKRADAKEQKPKGRAKGVPGEL 298
Qy 301 ATPDKKENDAKSSDSSVGEETLPSLSLPEKKVAEAEKKVEPAKKAEDQKEEDRRNYPT 360
Db 299 ATPDKKENDAKSSDSSVGEETLPSLSLPEKKVAEAEKKVEPAKKAEDQKEEDRRNYPT 358
Qy 361 NTYKTLLEIAESDVVEVKKAELVLKBEAKPRNEEKVKQAKAEVSKKAEATRLLEIKT 420
Db 359 NTYKTLLEIAESDVVEVKKAELVLKBEAKPRNEEKVKQAKAEVSKKAEATRLLEIKT 418
Qy 421 DRKKAEEBAKKAABEDKVKKEPAQPOPPAPAPAKAEKPAPAPKPNPAEQPKAEKPADOQ 480
Db 419 DRKKAEEBAKKAABEDKVKKEPAQPOPPAPAPAKAEKPAPAPKPNPAEQPKAEKPADOQ 478
Qy 481 AEEDYARRSEBEYNRLTQQQPPKTEKPAQSTP 513
Db 479 AEEDYARRSEBEYNRLTQQQPPKTEKPAQSTP 511

RESULT 6
US-10-341-201-13
; Sequence 13, Application US/10341201
; Publication No. US20050196405A1
; GENERAL INFORMATION:
```

; APPLICANT: BRILES et al.
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEIN C (PSPC), EPITOPIC REGIONS
; FILE REFERENCE: 454312-3140
; CURRENT APPLICATION NUMBER: US/10/341,201
; PRIOR FILING DATE: 2003-01-13
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 511
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-341-201-13

Query Match 98.8%; Score 2516; DB 18; Length 511;
Best Local Similarity 99.6%; Pred. No. 2.9e-106;
Matches 511; Conservative 0; Mismatches 0; Indels 2; Gaps 2;

Qy 1 MFASKSERKVVHYSIRKFSVGVASVVVASLVMSVGHATENEGATQVPTSSNRANESQAEQ 60
Db 1 MFASKSERKVVHYSIRKFSVGV-ASVVVASLVMSVGHATENEGATQVPTSSNRANESQAEQ 59

Qy 61 GEQPKLDSERDKARKEVEEYVKIVGESYAKSTKGRHTITVALVNLNINIKNEYLNKIV 120
Db 60 GEQPKLDSERDKARKEVEEYVKIVGESYAKSTKGRHTITVALVNLNINIKNEYLNKIV 119

Qy 121 ESTSESOLQILMMESRSKVDKAEVSKFEKSSSSSSSSSTKPEASDTAKPNKPTPEGEKV 180
Db 120 ESTSESOLQILMMESRSKVDKAEVSKFEKSSSSSSSSSTKPEASDTAKPNKPTPEGEKV 179

Qy 181 AEAKKVEEAEKKAQKQKEEDRRNYPTITYKTLEIAESDVEVKAELELVKVKANEPR 240
Db 180 AEAKKVEEA-KKAQKQKEEDRRNYPTITYKTLEIAESDVEVKAELELVKVKANEPR 238

Qy 241 DEQIKQAEAEVESKQAEATRLKKIKTDREAEAEAKRRADAKQKPKGRKRGVPGEL 300
Db 239 DEQIKQAEAEVESKQAEATRLKKIKTDREAEAEAKRRADAKQKPKGRKRGVPGEL 298

Qy 301 ATPDKENDAKSSDSSVGEETLPSPSLKPEKKVAEAEKVEEAKKAEADQKEDRRNYPT 360
Db 299 ATPDKENDAKSSDSSVGEETLPSPSLKPEKKVAEAEKVEEAKKAEADQKEDRRNYPT 358

Qy 361 NTYKTLLEIAESDVEVKAELELVKEEAKPRNEEKVKQAEVSKKAEATRLKIKT 420
Db 359 NTYKTLLEIAESDVEVKAELELVKEEAKPRNEEKVKQAEVSKKAEATRLKIKT 418

Qy 421 DRKKAEEAEKKAEEEDKVEKPAEQPAPAPKAEKPAPAPKAPENPAEQPKAEKPADQ 480
Db 419 DRKKAEEAEKKAEEEDKVEKPAEQPAPAPKAEKPAPAPKAPENPAEQPKAEKPADQ 478

Qy 481 AEEDYARRSEENRLTQQQPPKTEKPAQSTP 513
Db 479 AEEDYARRSEENRLTQQQPPKTEKPAQSTP 511

RESULT 7
US-09-748-875-12
; Sequence 12, Application US/09748875
; Publication No. US20010016200A1
; GENERAL INFORMATION:
; APPLICANT: BRILES et al.
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEIN C (PSPC), EPITOPIC REGIONS
; FILE REFERENCE: 454312-3140
; CURRENT APPLICATION NUMBER: US/09/748,875
; PRIOR FILING DATE: 2000-12-26
; PRIOR FILING DATE: 1999-04-23
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
; LENGTH: 513
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-341-201-13

; ORGANISM: Streptococcus pneumoniae
US-09-748-875-12

Query Match 98.4%; Score 2505; DB 9; Length 513;
Best Local Similarity 99.2%; Pred. No. 9.2e-106;
Matches 509; Conservative 1; Mismatches 1; Indels 2; Gaps 2;

Qy 1 MFASKSERKVVHYSIRKFSVGVASVVVASLVMSVGHATENEGATQVPTSSNRANESQAEQ 60
Db 1 MFASKSERKVVHYSIRKFSVGV-ASVVVASLVMSVGHATENEGATQVPTSSNRANESQAEQ 59

Qy 61 GEQPKLDSERDKARKEVEEYVKIVGESYAKSTKGRHTITVALVNLNINIKNEYLNKIV 120
Db 60 GEQPKLDSERDKARKEVEEYVKIVGESYAKSTKGRHTITVALVNLNINIKNEYLNKIV 119

Qy 121 ESTSESOLQILMMESRSKVDKAEVSKFEKSSSSSSSSSTKPEASDTAKPNKPTPEGEKV 180
Db 120 ESTSESOLQILMMESRSKVDKAEVSKFEKSSSSSSSSSTKPEASDTAKPNKPTPEGEKV 179

Qy 181 AEAKKVEEAEKKAQKQKEEDRRNYPTITYKTLEIAESDVEVKAELELVKVKANEPR 240
Db 180 AEAKKVEEA-KKAQKQKEEDRRNYPTITYKTLEIAESDVEVKAELELVKVKANEPR 238

Qy 241 DEQIKQAEAEVESKQAEATRLKKIKTDREAEAEAKRRADAKQKPKGRKRGVPGEL 300
Db 239 DEQIKQAEAEVESKQAEATRLKKIKTDREAEAEAKRRADAKQKPKGRKRGVPGEL 298

Qy 301 ATPDKENDAKSSDSSVGEETLPSPSLKPEKKVAEAEKVEEAKKAEADQKEDRRNYPT 360
Db 299 ATPDKENDAKSSDSSVGEETLPSPSLKPEKKVAEAEKVEEAKKAEADQKEDRRNYPT 358

Qy 361 NTYKTLLEIAESDVEVKAELELVKEEAKPRNEEKVKQAEVSKKAEATRLKIKT 420
Db 359 NTYKTLLEIAESDVEVKAELELVKEEAKPRNEEKVKQAEVSKKAEATRLKIKT 418

Qy 421 DRKKAEEAEKKAEEEDKVEKPAEQPAPAPKAEKPAPAPKAPENPAEQPKAEKPADQ 480
Db 419 DRKKAEEAEKKAEEEDKVEKPAEQPAPAPKAEKPAPAPKAPENPAEQPKAEKPADQ 478

Qy 481 AEEDYARRSEENRLTQQQPPKTEKPAQSTP 513
Db 479 AEEDYARRSEENRLTQQQPPKTEKPAQSTP 511

RESULT 8
US-09-298-523B-12
; Sequence 12, Application US/09298523B
; Publication No. US20030059438A1
; GENERAL INFORMATION:
; APPLICANT: BRILES et al.
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEIN C (PSPC), EPITOPIC REGIONS
; FILE REFERENCE: 454312-3140
; CURRENT APPLICATION NUMBER: US/09/298,523B
; PRIOR FILING DATE: 1999-04-23
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
; LENGTH: 513
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-298-523B-12

Query Match 98.4%; Score 2505; DB 10; Length 513;
Best Local Similarity 99.2%; Pred. No. 9.2e-106;
Matches 509; Conservative 1; Mismatches 1; Indels 2; Gaps 2;

Qy 1 MFASKSERKVVHYSIRKFSVGVASVVVASLVMSVGHATENEGATQVPTSSNRANESQAEQ 60
Db 1 MFASKSERKVVHYSIRKFSVGV-ASVVVASLVMSVGHATENEGATQVPTSSNRANESQAEQ 59

Qy 61 GEQPKLDSERDKARKEVEEYVKIVGESYAKSTKGRHTITVALVNLNINIKNEYLNKIV 120
Db 60 GEQPKLDSERDKARKEVEEYVKIVGESYAKSTKGRHTITVALVNLNINIKNEYLNKIV 119

Db 60 GEOPKLLDSERDKARKEVEEYVKIIVGESYAKSTKKRHTITVALVNLNINKNEYLNKIV 119
Qy 121 ESTSESQILMMESRSKVDVAVSKFEKXSSSSSSSSSTKPEASDTAKPNKPTSPGKV 180
Db 120 ESTSESQILMMESRSKVDVAVSKFEKXSSSSSSSSSTKPEASDTAKPNKPTSPGKV 179
Qy 181 AEAKKKVEAEAKKADQKEEDRRNYPITYKTLELEIAESDVEVKKAELELVKVKANEPR 240
Db 180 AEAKKKVEAEAKKADQKEEDRRNYPITYKTLELEIAESDVEVKKAELELVKVKANEPR 238
Qy 241 DEQIKIQAEEAVESKQAEATRLKKIKTDREAEAEAKRRADAKGQKPKGRAGKVGPEL 300
Db 239 DEQIKIQAEEAVESKQAEATRLKKIKTDREAEAEAKRRADAKGQKPKGRAGKVGPEL 298
Qy 301 ATPDKENDAKSSDSSVGEETLPSPLKPEKVAEAEKKVBEAKKABDQKEEDRRNYP 360
Db 299 ATPDKENDAKSSDSSVGEETLPSPLKPEKVAEAEKKVBEAKKABDQKEEDRRNYP 358
Qy 361 NTYKTLELEIAESDVEVKKAELELVKEEAKPRNEEKVKQAESVSKAEATRLKIKT 420
Db 359 NTYKTLELEIAESDVEVKKAELELVKEEAKPRNEEKVKQAESVSKAEATRLKIKT 418
Qy 421 DRKKAEEAEKKAEEEDKVKKEPAEQPOPAPAPKAEPAPAPKPNPAEQPKAEKPADQ 480
Db 419 DRKKAEEAEKKAEEEDKVKKEPAEQPOPAPAPKAEPAPAPKPNPAEQPKAEKPADQ 478
Qy 481 ABEYARRSEBYNRLTQQPPKTEKPAQPSTP 513
Db 479 ABEYARRSEBYNRLTQQPPKTEKPAQPSTP 511

RESULT 9

US-10-341-201-12
; Sequence 12, Application US/10341201
; Publication No. US20050196405A1
; GENERAL INFORMATION:
; APPLICANT: BRILES et al.
; TITLE OF INVENTION: PNEUMOCOCCAL SURFACE PROTEIN C (PSPC), EPITOPIC REGIONS
; FILE REFERENCE: 454312-3140
; CURRENT APPLICATION NUMBER: US/10/341,201
; CURRENT FILING DATE: 2003-01-13
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO 12
; LENGTH: 513
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-341-201-12

Query Match 98.4%; Score 2505; DB 18; Length 513;
Best Local Similarity 99.2%; Pred. No. 9.2e-106;
Matches 509; Conservative 1; Mismatches 1; Indels 2; Gaps 2;

Qy 1 MFASKSERKVHYSIRKFSVGVASVVVAVSLVMSVHVHATENEGATQVPTSSNRANESQAEQ 60
Db 1 MFASKSERKVHYSIRKFSVGV-ASVVAVSLVMSVHVHATENEGATQVPTSSNRANESQAEQ 59
Qy 61 GEOPKLLDSERDKARKEVEEYVKIIVGESYAKSTKKRHTITVALVNLNINKNEYLNKIV 120
Db 60 GEOPKLLDSERDKARKEVEEYVKIIVGESYAKSTKKRHTITVALVNLNINKNEYLNKIV 119
Qy 121 ESTSESQILMMESRSKVDVAVSKFEKXSSSSSSSSSTKPEASDTAKPNKPTSPGKV 180
Db 120 ESTSESQILMMESRSKVDVAVSKFEKXSSSSSSSSSTKPEASDTAKPNKPTSPGKV 179
Qy 181 AEAKKKVEAEAKKADQKEEDRRNYPITYKTLELEIAESDVEVKKAELELVKVKANEPR 240
Db 180 AEAKKKVEAEAKKADQKEEDRRNYPITYKTLELEIAESDVEVKKAELELVKVKANEPR 238
Qy 241 DEQIKIQAEEAVESKQAEATRLKKIKTDREAEAEAKRRADAKGQKPKGRAGKVGPEL 300
Db 239 DEQIKIQAEEAVESKQAEATRLKKIKTDREAEAEAKRRADAKGQKPKGRAGKVGPEL 298

Qy 301 ATPDKENDAKSSDSSVGEETLPSPLKPEKVAEAEKKVBEAKKABDQKEEDRRNYP 360
Db 299 ATPDKENDAKSSDSSVGEETLPSPLKPEKVAEAEKKVBEAKKABDQKEEDRRNYP 358
Qy 361 NTYKTLELEIAESDVEVKKAELELVKEEAKPRNEEKVKQAESVSKAEATRLKIKT 420
Db 359 NTYKTLELEIAESDVEVKKAELELVKEEAKPRNEEKVKQAESVSKAEATRLKIKT 418
Qy 421 DRKKAEEAEKKAEEEDKVKKEPAEQPOPAPAPKAEPAPAPKPNPAEQPKAEKPADQ 480
Db 419 DRKKAEEAEKKAEEEDKVKKEPAEQPOPAPAPKAEPAPAPKPNPAEQPKAEKPADQ 478
Qy 481 ABEYARRSEBYNRLTQQPPKTEKPAQPSTP 513
Db 479 ABEYARRSEBYNRLTQQPPKTEKPAQPSTP 511

RESULT 10

US-09-056-019-2
; Sequence 2, Application US/09056019A
; Publication No. US20030096950A1
; GENERAL INFORMATION:
; APPLICANT: Tuomanen, Elaine I
; APPLICANT: Wizemann, Theresa
; APPLICANT: Masure, H. R.
; APPLICANT: Johnson, Leslie S.
; TITLE OF INVENTION: POLYPEPTIDE COMPRISING THE AMINO ACID OF AN N-TERMINAL
; TITLE OF INVENTION: CHOLINE BINDING PROTEIN A TRUNCATE, VACCINE DERIVED
; TITLE OF INVENTION: THEREFROM AND USES THEREOF
; FILE REFERENCE: 1340-1-017
; CURRENT APPLICATION NUMBER: US/09/056.019A
; CURRENT FILING DATE: 1998-04-07
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 655
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-056-019-2

Query Match 92.9%; Score 2366; DB 10; Length 655;
Best Local Similarity 100.0%; Pred. No. 2.2e-99;
Matches 475; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 39 ENEGATQVPTSSNRANESQAEQGEQPKLLDSERDKARKEVEEYVKIIVGESYAKSTKKRH 98
Db 1 ENEGATQVPTSSNRANESQAEQGEQPKLLDSERDKARKEVEEYVKIIVGESYAKSTKKRH 60
Qy 99 TITVALVNLNINKNEYLNKIVESTSESQILMMESRSKYDEAVSKFEDSSSSSSSDS 158
Db 61 TITVALVNLNINKNEYLNKIVESTSESQILMMESRSKYDEAVSKFEDSSSSSSSDS 120
Qy 159 STKPEASDTAKPNKPTSPGKVAEAKKVAEAEKKADQKEEDRRNYPITYKTLELEIA 218
Db 121 STKPEASDTAKPNKPTSPGKVAEAKKVAEAEKKADQKEEDRRNYPITYKTLELEIA 180
Qy 219 ESDVEVKKAELELVKVKANEPRDEQIKIQAEEAVESKQAEATRLKKIKTDREAEAEAKR 278
Db 181 ESDVEVKKAELELVKVKANEPRDEQIKIQAEEAVESKQAEATRLKKIKTDREAEAEAKR 240
Qy 279 RADAKGQKPKGRAGKVGPELATPDKKENDAKSSDSSVGEETLPSPLKPEKPKVAEAK 338
Db 241 RADAKGQKPKGRAGKVGPELATPDKKENDAKSSDSSVGEETLPSPLKPEKPKVAEAK 300
Qy 339 KVBEAKKADQKEEDRRNYPITYKTLELEIAESDVEVKKAELELVKEEAKPRNEEKV 398
Db 301 KVBEAKKADQKEEDRRNYPITYKTLELEIAESDVEVKKAELELVKEEAKPRNEEKV 360
Qy 399 KQAEAVESKKAETRLKIKTDREAEAEAKRKAEEEDKVKKEPAEQPOPAPAPKAEPK 458
Db 361 KQAEAVESKKAETRLKIKTDREAEAEAKRKAEEEDKVKKEPAEQPOPAPAPKAEPK 420

```
Qy 459 APAPKPNPAQPKAEKPADQQAEDYARRSEEEYNRLTQQPPKTEKPAQPSPT 513
|||||
Db 421 APAPKPNPAQPKAEKPADQQAEDYARRSEEEYNRLTQQPPKTEKPAQPSPT 475

RESULT 11
US-10-751-702-2
; Sequence 2, Application US/10751702
; Publication No. US20040120966A1
; GENERAL INFORMATION:
; APPLICANT: Tuomanen, Elaine I
; APPLICANT: Wizemann, Theresa M.
; APPLICANT: Masure, H. R.
; APPLICANT: Johnson, Leslie S.
; APPLICANT: Koenig, Scott
; TITLE OF INVENTION: POLYPEPTIDE COMPRISING THE AMINO ACID OF AN N-TERMINAL
; TITLE OF INVENTION: CHOLINE BINDING PROTEIN A TRUNCATE, VACCINE DERIVED
; TITLE OF INVENTION: THEREFROM AND USES THEREOF
; FILE REFERENCE: 5853-2
; CURRENT APPLICATION NUMBER: US/10751,702
; PRIOR FILING DATE: 2004-01-05
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 655
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-751-702-2

Query Match 92.9%; Score 2366; DB 16; Length 655;
Best Local Similarity 100.0%; Pred. No. 2.2e-99;
Matches 475; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 39 ENEGATQVPTSSNRANESQAEQGGQPKLDSERDKARKEVEEYVKTI VGESYAKSTKKRH 98
|||||
Db 1 ENEGATQVPTSSNRANESQAEQGGQPKLDSERDKARKEVEEYVKTI VGESYAKSTKKRH 60

Qy 99 TITVALVNLNNIKNEYLKIVESTSESQILMMESRSKVDEAVSFKDSSSSSSSDS 158
|||||
Db 61 TITVALVNLNNIKNEYLKIVESTSESQILMMESRSKVDEAVSFKDSSSSSSSDS 120

Qy 159 STKPEASDTAKPNKPTBPGEKVAEAKKKVBEAEKKAQDQKEEDRRNPTTYKTLELEIA 218
|||||
Db 121 STKPEASDTAKPNKPTBPGEKVAEAKKKVBEAEKKAQDQKEEDRRNPTTYKTLELEIA 180

Qy 219 ESDVEVKKAELVVKVKAEPDRDQKI KQAEAEVESKQAEATRLKKIKTDREAEAEAKR 278
|||||
Db 181 ESDVEVKKAELVVKVKAEPDRDQKI KQAEAEVESKQAEATRLKKIKTDREAEAEAKR 240

Qy 279 RADAKEGQPKGRAKRGVPGELATPDKKENDAKSSDSVSGEETLPSPSLKPEKKVAEAEK 338
|||||
Db 241 RADAKEGQPKGRAKRGVPGELATPDKKENDAKSSDSVSGEETLPSPSLKPEKKVAEAEK 300

Qy 339 KVEEAKKAQDQKEEDRRNPTTYKTLELEIAESDVEVKKAELVVKERAKPRNEEKV 398
|||||
Db 301 KVEEAKKAQDQKEEDRRNPTTYKTLELEIAESDVEVKKAELVVKERAKPRNEEKV 360

Qy 399 KQAEAEVESKKAATRLKKIKTDREAEAEAKRKAEEEDVKKEKPAEQPAPAKAEKP 458
|||||
Db 361 KQAEAEVESKKAATRLKKIKTDREAEAEAKRKAEEEDVKKEKPAEQPAPAKAEKP 420

Qy 459 APAPKPNPAQPKAEKPADQQAEDYARRSEEEYNRLTQQPPKTEKPAQPSPT 513
|||||
Db 421 APAPKPNPAQPKAEKPADQQAEDYARRSEEEYNRLTQQPPKTEKPAQPSPT 475

RESULT 12
US-09-056-019-38
; Sequence 38, Application US/09056019A
; Publication No. US20030096950A1
```

```
; GENERAL INFORMATION:
; APPLICANT: Tuomanen, Elaine I
; APPLICANT: Wizemann, Theresa
; APPLICANT: Masure, H. R.
; APPLICANT: Johnson, Leslie S.
; APPLICANT: Koenig, Scott
; TITLE OF INVENTION: POLYPEPTIDE COMPRISING THE AMINO ACID OF AN N-TERMINAL
; TITLE OF INVENTION: CHOLINE BINDING PROTEIN A TRUNCATE, VACCINE DERIVED
; TITLE OF INVENTION: THEREFROM AND USES THEREOF
; FILE REFERENCE: 1340-1-017
; CURRENT APPLICATION NUMBER: US/09/056,019A
; CURRENT FILING DATE: 1998-04-07
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 38
; LENGTH: 460
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-056-019-38

Query Match 89.1%; Score 2268; DB 10; Length 460;
Best Local Similarity 99.8%; Pred. No. 4.2e-95;
Matches 458; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 25 VVASLVGMSVVHATENEGATQVPTSSNRANESQAEQGGQPKLDSERDKARKEVEEYVK 84
|||||
Db 2 IVASLVGMSVVHATENEGATQVPTSSNRANESQAEQGGQPKLDSERDKARKEVEEYVK 61

Qy 85 IVGESYAKSTKKGHTTITVALVNLNNIKNEYLKIVESTSESQILMMESRSKVDEAVS 144
|||||
Db 62 IVGESYAKSTKKGHTTITVALVNLNNIKNEYLKIVESTSESQILMMESRSKVDEAVS 121

Qy 145 KFEKDSSSSSSDSSTKPEASDTAKPNKPTBPGEKVAEAKKKVBEAEKKAQDQKEEDRRN 204
|||||
Db 122 KFEKDSSSSSSDSSTKPEASDTAKPNKPTBPGEKVAEAKKKVBEAEKKAQDQKEEDRRN 181

Qy 205 YPTITTYKTLELEIAESDVEVKKAELVVKVKAEPDRDQKI KQAEAEVESKQAEATRLKK 264
|||||
Db 182 YPTITTYKTLELEIAESDVEVKKAELVVKVKAEPDRDQKI KQAEAEVESKQAEATRLKK 241

Qy 265 IKTDREAEAEAKRADAKEGQPKGRAKRGVPGELATPDKKENDAKSSDSVSGEETLPS 324
|||||
Db 242 IKTDREAEAEAKRADAKEGQPKGRAKRGVPGELATPDKKENDAKSSDSVSGEETLPS 301

Qy 325 PSLLKPEKKVAEAEKKEEAKKKAEDQKEEDRRNPTTYKTLELEIAESDVEVKKAEL 384
|||||
Db 302 PSLLKPEKKVAEAEKKEEAKKKAEDQKEEDRRNPTTYKTLELEIAESDVEVKKAEL 361

Qy 385 VKEEAKEPRNEEKVKQAEVESKKAATRLKKIKTDREAEAEAKRKAEEEDVKKEKPA 444
|||||
Db 362 VKEEAKEPRNEEKVKQAEVESKKAATRLKKIKTDREAEAEAKRKAEEEDVKKEKPA 421

Qy 445 EQQPAPAPAEKPAKPAKPNPAEQPKAEKPADQQAEE 483
|||||
Db 422 EQQPAPAPAEKPAKPAKPNPAEQPKAEKPADQQAEE 460

RESULT 13
US-10-751-702-38
; Sequence 38, Application US/10751702
; Publication No. US20040120966A1
; GENERAL INFORMATION:
; APPLICANT: Tuomanen, Elaine I
; APPLICANT: Wizemann, Theresa M.
; APPLICANT: Masure, H. R.
; APPLICANT: Johnson, Leslie S.
; APPLICANT: Koenig, Scott
; TITLE OF INVENTION: POLYPEPTIDE COMPRISING THE AMINO ACID OF AN N-TERMINAL
; TITLE OF INVENTION: CHOLINE BINDING PROTEIN A TRUNCATE, VACCINE DERIVED
; TITLE OF INVENTION: THEREFROM AND USES THEREOF
; FILE REFERENCE: 5853-2
; CURRENT APPLICATION NUMBER: US/10751,702
; CURRENT FILING DATE: 2004-01-05
```

```
; PRIOR APPLICATION NUMBER: US/09/056,019B
; PRIOR FILING DATE: 1998-04-07
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 38
; LENGTH: 460
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-751-702-38

Query Match      89.1%; Score 2268; DB 16; Length 460;
Best Local Similarity 99.8%; Pred. No. 4.2e-95;
Matches 458; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 25 VVASLVMSGVVHATENEGATQVPTSSNRANESQAEGQPKKLDSEDKARKEVEEYVKK 84
   :|||||
DB 2 IVASLVMSGVVHATENEGATQVPTSSNRANESQAEGQPKKLDSEDKARKEVEEYVKK 61
   :|||||
QY 85 IVGESYAKSTKKRHTITVALVNLNINIKNEYLNKIVESTSESQLOILMMESRSKVDEAVS 144
   :|||||
DB 61 IVGESYAKSTKKRHTITVALVNLNINIKNEYLNKIVESTSESQLOILMMESRSKVDEAVS 120
   :|||||
QY 145 KFEKDSSTSSSSDSTKPEASDTAKPNKPTPEGKVAEAKKVEAEKKAQKQKEEDRRN 204
   :|||||
DB 121 KFEKDSSTSSSSDSTKPEASDTAKPNKPTPEGKVAEAKKVEAEKKAQKQKEEDRRN 180
   :|||||
QY 205 YPTITYKTLELEIAESDVEVKKAELVVKVANEPRDEQIKIQAABEVEKKAQAEATRLKK 264
   :|||||
DB 181 YPTITYKTLELEIAESDVEVKKAELVVKVANEPRDEQIKIQAABEVEKKAQAEATRLKK 240
   :|||||
QY 265 IKTDREAEAEERAKRADAKEQKPKGRKRGVPGELATPDKKENDAKSSDSVSGEETLPS 324
   :|||||
DB 241 IKTDREAEAEERAKRADAKEQKPKGRKRGVPGELATPDKKENDAKSSDSVSGEETLPS 300
   :|||||
QY 325 PSLKPEKKVAEAEKKVEAEKKAQKQKEEDRRNYPNTYKTLELEIAESDVEVKKAEL 384
   :|||||
DB 301 PSLKPEKKVAEAEKKVEAEKKAQKQKEEDRRNYPNTYKTLELEIAESDVEVKKAEL 360
   :|||||
QY 385 VKEEAKPRNEBEKVQAQAEVSKAAEATRLBKIKTDRKKAEEAEAKRKAABEDKVKEKPA 444
   :|||||
DB 361 VKEEAKPRNEBEKVQAQAEVSKAAEATRLBKIKTDRKKAEEAEAKRKAABEDKVKEKPA 420
   :|||||
QY 445 EQQPAPAPKAPKAPAPKAPENPAPQPKAEKPADQQAEE 483
   :|||||
DB 421 EQQPAPAPKAPKAPAPKAPENPAPQPKAEKPADQQAEE 459
   :|||||

RESULT 15
US-10-751-702-39
; Sequence 39, Application US/10751702
; Publication No. US20040120966A1
; GENERAL INFORMATION:
; APPLICANT: Tuomanen, Elaine I
; APPLICANT: Wisemann, Theresa M.
; APPLICANT: Masure, H. R.
; APPLICANT: Johnson, Leslie S.
; TITLE OF INVENTION: POLYPEPTIDE COMPRISING THE AMINO ACID OF AN N-TERMINAL
; TITLE OF INVENTION: CHOLINE BINDING PROTEIN A TRUNCATE, VACCINE DERIVED
; TITLE OF INVENTION: THEREFROM AND USES THEREOF
; FILE REFERENCE: 5853-2
; CURRENT APPLICATION NUMBER: US/10/751,702
; CURRENT FILING DATE: 2004-01-05
; PRIOR APPLICATION NUMBER: US/09/056,019B
; PRIOR FILING DATE: 1998-04-07
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 39
; LENGTH: 459
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-751-702-39

Query Match      88.4%; Score 2251; DB 16; Length 459;
Best Local Similarity 98.9%; Pred. No. 2.5e-94;
Matches 454; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 25 VVASLVMSGVVHATENEGATQVPTSSNRANESQAEGQPKKLDSEDKARKEVEEYVKK 84
   :|||||
DB 1 IVASLVMSGVVHATENEGATQVPTSSNRANESQAEGQPKKLDSEDKARKEVEEYVKK 60
   :|||||
QY 85 IVGESYAKSTKKRHTITVALVNLNINIKNEYLNKIVESTSESQLOILMMESRSKVDEAVS 144
   :|||||
DB 61 IVGESYAKSTKKRHTITVALVNLNINIKNEYLNKIVESTSESQLOILMMESRSKVDEAVS 120
   :|||||

Query Match      88.4%; Score 2251; DB 10; Length 459;
Best Local Similarity 98.9%; Pred. No. 2.5e-94;

; PRIOR APPLICATION NUMBER: US/09056019A
; PRIOR FILING DATE: 1998-04-07
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 38
; LENGTH: 460
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-056-019-39

Query Match      89.1%; Score 2268; DB 16; Length 460;
Best Local Similarity 99.8%; Pred. No. 4.2e-95;
Matches 458; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 25 VVASLVMSGVVHATENEGATQVPTSSNRANESQAEGQPKKLDSEDKARKEVEEYVKK 84
   :|||||
DB 2 IVASLVMSGVVHATENEGATQVPTSSNRANESQAEGQPKKLDSEDKARKEVEEYVKK 61
   :|||||
QY 85 IVGESYAKSTKKRHTITVALVNLNINIKNEYLNKIVESTSESQLOILMMESRSKVDEAVS 144
   :|||||
DB 61 IVGESYAKSTKKRHTITVALVNLNINIKNEYLNKIVESTSESQLOILMMESRSKVDEAVS 121
   :|||||
QY 145 KFEKDSSTSSSSDSTKPEASDTAKPNKPTPEGKVAEAKKVEAEKKAQKQKEEDRRN 204
   :|||||
DB 122 KFEKDSSTSSSSDSTKPEASDTAKPNKPTPEGKVAEAKKVEAEKKAQKQKEEDRRN 181
   :|||||
QY 205 YPTITYKTLELEIAESDVEVKKAELVVKVANEPRDEQIKIQAABEVEKKAQAEATRLKK 264
   :|||||
DB 182 YPTITYKTLELEIAESDVEVKKAELVVKVANEPRDEQIKIQAABEVEKKAQAEATRLKK 241
   :|||||
QY 265 IKTDREAEAEERAKRADAKEQKPKGRKRGVPGELATPDKKENDAKSSDSVSGEETLPS 324
   :|||||
DB 242 IKTDREAEAEERAKRADAKEQKPKGRKRGVPGELATPDKKENDAKSSDSVSGEETLPS 301
   :|||||
QY 325 PSLKPEKKVAEAEKKVEAEKKAQKQKEEDRRNYPNTYKTLELEIAESDVEVKKAEL 384
   :|||||
DB 302 PSLKPEKKVAEAEKKVEAEKKAQKQKEEDRRNYPNTYKTLELEIAESDVEVKKAEL 361
   :|||||
QY 385 VKEEAKPRNEBEKVQAQAEVSKAAEATRLBKIKTDRKKAEEAEAKRKAABEDKVKEKPA 444
   :|||||
DB 362 VKEEAKPRNEBEKVQAQAEVSKAAEATRLBKIKTDRKKAEEAEAKRKAABEDKVKEKPA 421
   :|||||
QY 445 EQQPAPAPKAPKAPAPKAPENPAPQPKAEKPADQQAEE 483
   :|||||
DB 422 EQQPAPAPKAPKAPAPKAPENPAPQPKAEKPADQQAEE 460
   :|||||

RESULT 14
US-09-056-019-39
; Sequence 39, Application US/09056019A
; Publication No. US20030096950A1
; GENERAL INFORMATION:
; APPLICANT: Tuomanen, Elaine I
; APPLICANT: Wisemann, Theresa
; APPLICANT: Masure, H. R.
; APPLICANT: Johnson, Leslie S.
; TITLE OF INVENTION: POLYPEPTIDE COMPRISING THE AMINO ACID OF AN N-TERMINAL
; TITLE OF INVENTION: CHOLINE BINDING PROTEIN A TRUNCATE, VACCINE DERIVED
; TITLE OF INVENTION: THEREFROM AND USES THEREOF
; FILE REFERENCE: 1340-1-017
; CURRENT APPLICATION NUMBER: US/09/056,019A
; CURRENT FILING DATE: 1998-04-07
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 39
; LENGTH: 459
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-056-019-39
```

QY	145	KPEKDSSSSSDSTKPEASDTAKPNKPTPEGEKVAEAKKKVEEAEKKAQKQKEEDRRN	204
Db	121		
QY	205	YPTIYKTLLELEIAESDVEVKAELELVKVKANEPRDQKIQAEEVESKQAEATRLLK	264
Db	181		
QY	265	IKTDREAEAEAKRRADAKEQKPKGRKRGVPGELATPDKKENDAKSSDSSVGEETLPS	324
Db	241		
QY	325	PSLKPEKKVAEAKKVEEAKKAEDQKBEDRRNPTNTYKTLLELEIAESDVEVKAELEL	384
Db	301		
QY	385	VKEEAKEPNNEKVKQAKAEVESKKAETRLKIKTDRKKAEEAKKKAABEDVKKEKPA	444
Db	361		
QY	445	EQQPAPAPKAEKPAPAPKPNPAPAEQPKAEKPADOQAAE	483
Db	421	EQQPAPAPKTEKPAPAPKPNPAPAEQPKAEKPADOQAAE	459

Search completed: September 28, 2005, 18:56:50
Job time : 176 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 28, 2005, 18:33:49 ; Search time 43 Seconds
(without alignments)
890.581 Million cell updates/sec

Title: US-09-298-523D-13
Perfect score: 2546
Sequence: 1 MFASKSRKVHYSIRKFSVG.....NRLTQQQPPKTEKPAQSTP 513

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA.*
1: /cgn2_6/ptodata/1/1aa/5A.COMB.pdp:*
2: /cgn2_6/ptodata/1/1aa/5B.COMB.pdp:*
3: /cgn2_6/ptodata/1/1aa/6A.COMB.pdp:*
4: /cgn2_6/ptodata/1/1aa/6B.COMB.pdp:*
5: /cgn2_6/ptodata/1/1aa/PTUS.COMB.pdp:*
6: /cgn2_6/ptodata/1/1aa/backfiles1.pdp:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2210	86.8	446	4	US-09-286-981B-9
2	2193	86.1	446	4	US-09-286-981B-6
3	1754	68.9	564	3	US-09-308-022-6
4	1524	59.9	406	4	US-09-286-981B-18
5	1517.5	59.6	631	3	US-08-847-065-25
6	1517.5	59.6	631	4	US-09-829-382-25
7	1467	57.6	426	4	US-09-286-981B-12
8	1459	57.3	425	4	US-09-286-981B-13
9	1458	57.3	424	4	US-09-286-981B-14
10	1452	57.0	428	4	US-09-286-981B-7
11	1443.5	56.7	425	4	US-09-286-981B-11
12	1428	56.1	412	4	US-09-286-981B-17
13	1417	55.7	414	4	US-09-286-981B-16
14	1416	55.6	419	4	US-09-286-981B-15
15	1413	55.5	414	4	US-09-286-981B-10
16	1410.5	55.4	431	4	US-09-286-981B-3
17	1373	53.9	453	3	US-08-961-083-38
18	1373	53.9	453	4	US-09-536-784-38
19	1306.5	51.3	413	4	US-09-286-981B-5
20	1233	48.4	864	4	US-08-714-741-40
21	1168.5	45.9	588	4	US-08-714-741-42
22	915.5	36.0	605	4	US-08-714-741-46
23	904.5	35.5	222	4	US-09-107-433-4344
24	879.5	34.5	216	4	US-09-583-110-4344
25	854	33.5	8991	4	US-08-714-741-32
26	775.5	30.5	1231	4	US-08-714-741-41
27	770	30.2	275	4	US-09-583-110-4347

28	757.5	29.8	623	4	US-08-714-741-47	Sequence 47, Appl
29	714	28.0	251	4	US-09-286-981B-4	Sequence 4, Appli
30	657	25.8	233	4	US-09-107-433-3526	Sequence 3526, Ap
31	636	25.0	142	3	US-08-847-065-21	Sequence 21, Appl
32	636	25.0	142	4	US-09-829-382-21	Sequence 21, Appl
33	595	23.4	131	4	US-08-529-055-56	Sequence 56, Appl
34	576.5	22.6	128	4	US-08-529-055-57	Sequence 57, Appl
35	561.5	22.1	482	4	US-09-107-433-4705	Sequence 4705, Ap
36	560	22.0	219	4	US-09-286-981B-8	Sequence 8, Appli
37	542	21.3	114	4	US-09-286-981B-19	Sequence 19, Appl
38	541	21.2	110	3	US-08-961-083-102	Sequence 102, App
39	541	21.2	110	4	US-09-536-784-102	Sequence 102, App
40	497	19.5	695	1	US-08-127-499A-23	Sequence 23, Appl
41	497	19.5	695	1	US-08-482-847-23	Sequence 23, Appl
42	496	19.5	619	1	US-08-465-746-2	Sequence 2, Appli
43	496	19.5	619	1	US-08-214-164-2	Sequence 2, Appli
44	496	19.5	619	2	US-08-467-852A-3	Sequence 3, Appli
45	496	19.5	619	2	US-08-246-636-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1

US-09-286-981B-9
; Sequence 9, Application US/09286981B
; Patent No. 6503511
; GENERAL INFORMATION:
; APPLICANT: Wizemann, Theresa M.
; APPLICANT: Koenig, Scott
; APPLICANT: Johnson, Leslie S
; TITLE OF INVENTION: Derivatives of Choline Binding Proteins for Vaccines
; FILE REFERENCE: 469201-396
; CURRENT APPLICATION NUMBER: US/09/286,981B
; CURRENT FILING DATE: 1999-04-06
; PRIOR APPLICATION NUMBER: US 60/085,743
; PRIOR FILING DATE: 1998-05-15
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 446
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-286-981B-9

Query Match	86.8%	Score	2210;	DB	4;	Length	446;
Best Local Similarity	100.0%	Pred. No.	2.9e-145;				
Matches	446;	Conservative	0;	Mismatches	0;	Indels	0;
Gaps	0;						
QY	38	TENEGATQVPTSSNRANESQAEGQEQPKLDSERDKARKEVEEYVKIVGESYAKSTKCR	97				
Db	1	TENEGATQVPTSSNRANESQAEGQEQPKLDSERDKARKEVEEYVKIVGESYAKSTKCR	60				
QY	98	HTITVALNELNNIKNEYLKIVESTSSQILMMESRSKVDEAVSKFEKDSSSSSSD	157				
Db	61	HTITVALNELNNIKNEYLKIVESTSSQILMMESRSKVDEAVSKFEKDSSSSSSD	120				
QY	158	SSTKPEASDTAKPNKPTPEGKVAEAKKKVEAEAKKADQKEEDRRNPTTYTKTLELEI	217				
Db	121	SSTKPEASDTAKPNKPTPEGKVAEAKKKVEAEAKKADQKEEDRRNPTTYTKTLELEI	180				
QY	218	AESDVEVKAELELVKVKANPRDQKIQAFAEVESQAEATRLKKIKTDREAEAEAK	277				
Db	181	AESDVEVKAELELVKVKANPRDQKIQAFAEVESQAEATRLKKIKTDREAEAEAK	240				
QY	278	PRADAKEGKPKRAKRGVPGELATPDKKENDAKSSDSSVGEETLPSLKEPKKVAEAE	337				
Db	241	PRADAKEGKPKRAKRGVPGELATPDKKENDAKSSDSSVGEETLPSLKEPKKVAEAE	300				
QY	338	KKVEAKKKAEDQKEEDRRNPTTYTKTLELEIAESDVEVKAELELVKKEAEKPRNEEK	397				
Db	301	KKVEAKKKAEDQKEEDRRNPTTYTKTLELEIAESDVEVKAELELVKKEAEKPRNEEK	360				

Db 259 KRAKRGVPGELATPDKKENDAKSSDSSVGBETLPSLSKSGKVAEAKKKVEAEKAK 318
Qy 349 DQKEEDRRNYPTNTYKTLLEIAESDVEVKKAELELVKEEAKPRNEEKVKQAKAEVSK 408
Db 319 DQKEEDRRNYPTNTYKTLLEIAESDVEVKKAELELVKEEAKPRNEEKVKQAKAEVSK 378
Qy 409 KAEATRELEKIITDRKAEAEAKKAAEEDVKKEKPAEQOPAPAPKAEKPAKAPENPA 468
Db 379 KAEATRELEKIITDRKAEAEAKKAAEEDVKKEKPAEQOPAPAPKAEKPAKAPENPA 436
Qy 469 EQPKAEKADQQAEDYARRSEEEYNRLTQOOPPKTEKPAQSTP 513
Db 437 EQPKAEKADQQAEDYARRSEEEYNRLTQOOPPKTEKPAQSTP 481

RESULT 4
US-09-286-981B-18
; Sequence 18, Application US/09286981B
; Patent No. 6503511

GENERAL INFORMATION:
; APPLICANT: Wismann, Theresa M.
; APPLICANT: Koenig, Scott
; APPLICANT: Johnson, Leslie S
; TITLE OF INVENTION: Derivatives of Choline Binding Proteins for Vaccines
; FILE REFERENCE: 469201-396
; CURRENT APPLICATION NUMBER: US/09/286,981B
; PRIOR FILING DATE: 1999-04-06
; PRIOR APPLICATION NUMBER: US 60/085,743
; PRIOR FILING DATE: 1998-05-15
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
; LENGTH: 406
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-286-981B-18

Query Match 59.9%; Score 1524; DB 4; Length 406;
Best Local Similarity 73.0%; Pred. No. 7.6e-98;
Matches 327; Conservative 29; Mismatches 46; Indels 46; Gaps 7;

Qy 38 TENEGATQVPTSSNRANESQAEQEQPKLSDERDKARKEVEEYVKKIVGESYAKTKR 97
Db 1 TENEGTQAPTSSNRGNESQA-----EHMKAAQVDEYIERML-----QLDRK 44
Qy 98 HTITVALVNELNINIKNEYLNKIVESTSESOLQILMMESRSKVDEAVSKFEKDSSSSSSD 157
Db 45 HTQNVGLLTGLGAKTEYLRGLSVSKSTAE-LPSEIKELTAAPKQFK----- 94
Qy 158 SSTPEASDTAKPNKPTPEGKVAEAKKVAEAKKAKQKQKEEDRRNYPTTYKTLLEI 217
Db 95 -----DTLKPEK-----KVAEAEKKVAEAKKAKQKQKEEDRRNYPTTYKTLLEI 140
Qy 218 AESDVEVKKAELELVKVKANPREDEOKIKQAEAVESKQAEATRLKIKTDRREAESEAK 277
Db 141 AESDVEVKKAELELVKVKANPREDEOKIKQAEAVESKQAEATRLKIKTDRREAESEAK 200
Qy 278 RRADAKEQ-----KPKGRKRGVPGELATPDKKENDAKSSDSSVGEETLPSPLKPKKV 333
Db 201 RRADAKEQDESSKRRKSRVKGDLGEQATPDKKENDAKSSDSSVGEETLPSPLKPKKV 260
Qy 334 AEAEKKVEEAKKADQKQEDRRNYPTNTYKTLLEIAESDVEVKKAELELVKEEAKPR 393
Db 261 AEAEKKVEEADKKAQKQEDRRNYPTNTYKTLLEIAESDVEVKKAELELVKEEAKPR 320
Qy 394 NEEKVKQAKAEVSKKAEATLEKIKTDRKAEAEKKAEEEDVKKEKPAEQOPAPAP 453
Db 321 NEEKVKQAKAEVSKKAEATLEKIKTDRKAEAEKKAEEEDVKKEKPAEQOPAPAP 380
Qy 454 KAEKPAKAPENPAEQPKAEKPADQQA 481
Db 381 QPEK--PAPKPENPAEQPKAEKPADQQA 406

RESULT 5

US-08-847-065-25
; Sequence 25, Application US/08847065
; Patent No. 6245335
GENERAL INFORMATION:
; APPLICANT: Masure, H. Robert
; APPLICANT: Rosenow, Carsten I.
; APPLICANT: Tuomanen, Elaine
; APPLICANT: Wismann, Theresa M.
; TITLE OF INVENTION: CHOLINE BINDING PROTEINS FOR
; TITLE OF INVENTION: ANTI-PNEUMOCOCCAL VACCINES
; NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
; ADDRESSEE: David A. Jackson, Esq.
; STREET: 411 Hackensack Ave, Continental Plaza, 4th
; STREET: Floor
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/847,065
; FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-158 ..
TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-487-5800
; TELEFAX: 201-343-1684

INFORMATION FOR SEQ ID NO: 25:
SEQUENCE CHARACTERISTICS:
; LENGTH: 631 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEICAL: NO
; FRAGMENT TYPE: N-terminal
US-08-847-065-25

Query Match 59.6%; Score 1517.5; DB 3; Length 631;
Best Local Similarity 73.5%; Pred. No. 3.6e-97;
Matches 316; Conservative 40; Mismatches 35; Indels 39; Gaps 5;

Qy 95 KCRHTITVALVNELNINIKNEYLNK--IVESTSESOLQILMMESRSKVDEAVSKFEKDS 152
Db 9 RRKHTQNVALNIKLSAKTKYLRNLNVLVEEKSQDEL---SEIKAKLDAAEKFKQDT-- 63
Qy 153 SSSSDSSTKPEASDTAKPNKPTPEGKVAEAKKVAEAKKAKQKQKEEDRRNYPTTYKT 212
Db 64 -----LKPGEKVAEAKKVAEAKKAKQKQKEEDRRNYPTNTYKT 102
Qy 213 LELEIAESDVEVKKAELELVKVKANPREDEOKIKQAEAVESKQAEATRLKIKTDRREA 272
Db 103 LELEIAFDVKVKEAELELVKEEAKESRNEGTIKQAEKVESKAEATRLNLIKTDKKA 162
Qy 273 EEEAKRADAK-----EQQPKGRKRGVPGELATPDKKENDAKSSDSSVGEETLP 323
Db 163 EEEAKRADAKLKEANVATSDQPKGRKRGVPGELATPDKKENDAKSSDSSVGEETLP 222
Qy 324 SPSLKPKKVAEAEKKVEEAKKADQKQEDRRNYPTNTYKTLLEIAESDVEVKKAELE 383
Db 223 SSSLKSKKVAEAEKKVEEAKKAKQKQEDRRNYPTNTYKTLLEIAESDVKVKEAELE 282
Qy 384 LVKEEAKPRNEEKVKQAKAEVSKKAEATRLKIKTDRKAEAEKKAEEEDVKKEK 443

QY 329 PEKKVAEAEKKVVEAKKAEADQKEDRRNYPNTYKTLLEIAESDVEVKAELELVKEE 388
DB 262 PEKKVAEAEKKVVEAKKAEADQKEDRRNYPNTYKTLLEIAESDVEVKAELELVKEE 321
QY 389 AKEPRNEEKVQKAEVSKAEATRLKIKTDKKA-BEEAKKAAEEDVKKEKPAQP 447
DB 322 AKESRNEEKIKQVKAKVESKAEATRLKIKTDKKA-BEEAKKAAEEDVKKEKPAQP 381
QY 448 QPAPAPKAEK-----PAPAPKPNPAPAPKPAEKPADQQAEE 483
DB 382 QPAPAPQEPKTEEPENPAPAPKPNPAPKPAEKPADQQAEE 426

RESULT 8
US-09-286-981B-13
; Sequence 13, Application US/09286981B
; Patent No. 6503511
; GENERAL INFORMATION:
; APPLICANT: Wizemann, Theresa M.
; APPLICANT: Koenig, Scott
; APPLICANT: Johnson, Leslie S
; TITLE OF INVENTION: Derivatives of Choline Binding Proteins for Vaccines
; FILE REFERENCE: 469201-396
; CURRENT APPLICATION NUMBER: US/09/286,981B
; PRIOR FILING DATE: 1999-04-06
; PRIOR APPLICATION NUMBER: US 60/085,743
; PRIOR FILING DATE: 1998-05-15
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 425
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-286-981B-13

Query Match 57.3%; Score 1459; DB 4; Length 425;
Best Local Similarity 68.3%; Pred. No. 2.5e-93;
Matches 317; Conservative 42; Mismatches 47; Indels 58; Gaps 8;

QY 38 TENEGATQVPTSSNRANESQAEQGEQPKLDSERDKARKEVEYVKIVGESYAKSTKKR 97
DB 1 TEKEVTTQVATSSNRANKSQ-----TEHMKAAKQVDEYIKKKL-----QLDRRK 44
QY 98 HTITVALVNLNINIKVEYLKIVESTSSQILMMESRSKVDKAVSKFEDSSSSSSD 157
DB 45 HTQNVGLLTGLGVIKTEYLHGLSVSKKSEAE-LPSEIKAKLDAAFEQFKKDT 96
QY 158 SSTPEASDTAKPNKPTPEGEKVAEAKKVEAEKKAQKQEDRRNYPNTYKTLLEI 217
DB 97 -----LPTPEGKKVAEAEKKVVEAKKAEADQKEDRRNYPNTYKTLLEI 142
QY 218 AESDVEVKAELELVKVKANEPDRQKIKQAEAEVSKAEATRLKIKTDREAEAEBAK 277
DB 143 AESDVEVKAELELVKVEAKESRDEKKNQAKAVENKKAATRLKIKTDREKA-BEAK 201
QY 278 RRADAK-----EQKPKGRAKGVPGELATPDKKENDAKSSDSSVGBETLPSPLK 328
DB 202 RRADAKQEAENVATSEODKSKRAKRVFGLATPDKKENDAKSSDSSVGBETLPSPLK 261
QY 329 PEKKVAEAEKKVVEAKKAEADQKEDRRNYPNTYKTLLEIAESDVEVKAELELVKEE 388
DB 262 PEKKVAEAEKKVVEAKKAEADQKEDRRNYPNTYKTLLEIAESDVEVKAELELVKEE 321
QY 389 AKEPRNEEKVQKAEVSKAEATRLKIKTDKKA-BEEAKKAAEEDVKKEKPAQP 447
DB 322 AKESRNEEKIKQVKAKVESKAEATRLKIKTDKKA-BEEAKKAAEEDVKKEKPAQP 381
QY 448 QPAPAPKAEK-----PAPAPKPNPAPAPKPAEKPADQQAEE 482
DB 382 QPAPAPQEPKTEEPENPAPAPKPNPAPKPAEKPADQQAEE 425

RESULT 9

US-09-286-981B-14
; Sequence 14, Application US/09286981B
; Patent No. 6503511
; GENERAL INFORMATION:
; APPLICANT: Wizemann, Theresa M.
; APPLICANT: Koenig, Scott
; APPLICANT: Johnson, Leslie S
; TITLE OF INVENTION: Derivatives of Choline Binding Proteins for Vaccines
; FILE REFERENCE: 469201-396
; CURRENT APPLICATION NUMBER: US/09/286,981B
; PRIOR FILING DATE: 1999-04-06
; PRIOR APPLICATION NUMBER: US 60/085,743
; PRIOR FILING DATE: 1998-05-15
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 14
; LENGTH: 424
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-286-981B-14

Query Match 57.3%; Score 1458; DB 4; Length 424;
Best Local Similarity 68.5%; Pred. No. 2.9e-93;
Matches 317; Conservative 41; Mismatches 47; Indels 58; Gaps 8;

QY 38 TENEGATQVPTSSNRANESQAEQGEQPKLDSERDKARKEVEYVKIVGESYAKSTKKR 97
DB 1 TEKEVTTQVATSSNRANKSQ-----TEHMKAAKQVDEYIKKKL-----QLDRRK 44
QY 98 HTITVALVNLNINIKVEYLKIVESTSSQILMMESRSKVDKAVSKFEDSSSSSSD 157
DB 45 HTQNVGLLTGLGVIKTEYLHGLSVSKKSEAE-LPSEIKAKLDAAFEQFKKDT 96
QY 158 SSTPEASDTAKPNKPTPEGEKVAEAKKVEAEKKAQKQEDRRNYPNTYKTLLEI 217
DB 97 -----LPTPEGKKVAEAEKKVVEAKKAEADQKEDRRNYPNTYKTLLEI 142
QY 218 AESDVEVKAELELVKVKANEPDRQKIKQAEAEVSKAEATRLKIKTDREAEAEBAK 277
DB 143 AESDVEVKAELELVKVEAKESRDEKKNQAKAVENKKAATRLKIKTDREKA-BEAK 201
QY 278 RRADAK-----EQKPKGRAKGVPGELATPDKKENDAKSSDSSVGBETLPSPLK 328
DB 202 RRADAKQEAENVATSEODKSKRAKRVFGLATPDKKENDAKSSDSSVGBETLPSPLK 261
QY 329 PEKKVAEAEKKVVEAKKAEADQKEDRRNYPNTYKTLLEIAESDVEVKAELELVKEE 388
DB 262 PEKKVAEAEKKVVEAKKAEADQKEDRRNYPNTYKTLLEIAESDVEVKAELELVKEE 321
QY 389 AKEPRNEEKVQKAEVSKAEATRLKIKTDKKA-BEEAKKAAEEDVKKEKPAQP 447
DB 322 AKESRNEEKIKQVKAKVESKAEATRLKIKTDKKA-BEEAKKAAEEDVKKEKPAQP 381
QY 448 QPAPAPKAEK-----PAPAPKPNPAPAPKPAEKPADQQAEE 481
DB 382 QPAPAPQEPKTEEPENPAPAPKPNPAPKPAEKPADQQAEE 424

RESULT 10
US-09-286-981B-7
; Sequence 7, Application US/09286981B
; Patent No. 6503511
; GENERAL INFORMATION:
; APPLICANT: Wizemann, Theresa M.
; APPLICANT: Koenig, Scott
; APPLICANT: Johnson, Leslie S
; TITLE OF INVENTION: Derivatives of Choline Binding Proteins for Vaccines
; FILE REFERENCE: 469201-396
; CURRENT APPLICATION NUMBER: US/09/286,981B
; PRIOR FILING DATE: 1999-04-06
; PRIOR APPLICATION NUMBER: US 60/085,743
; PRIOR FILING DATE: 1998-05-15
; NUMBER OF SEQ ID NOS: 38


```
Db 225 DFGPATPKDKENDAKSSSSVGEETLPSLUKPGKVAEAEKKVVEAEKKAAQKQKEDH 284
Qy 356 RNYPTNTYKTLLEIAESDVEVKAELELVKEEAEKEPRNEEKVKQAKAEVSKAEATRL 415
Db 285 RNYPTITYKTLLEIAESDVEVKAELELVKEEAGSRNEEKVKQAKAEVSKAEATRL 344
Qy 416 EKIKTRDKAEAEAKRAAEEDVKVEKPAEQPAPAPAKAEKPAAPAPKPPENPAEQPKAEK 475
Db 345 EKIKTRDKAEAEAKRAAEEDVKVEKPAEQPAPAPAKAEKPAAPAPKPPENPAEQPKAEK 404
Qy 476 PADQQAEE 483
Db 405 PADQQAEE 412

RESULT 13
US-09-286-981B-16
; Sequence 16, Application US/09286981B
; Patent No. 6503511
; GENERAL INFORMATION:
; APPLICANT: Wizemann, Theresa M.
; APPLICANT: Koenig, Scott
; APPLICANT: Johnson, Leslie S
; TITLE OF INVENTION: Derivatives of Choline Binding Proteins for Vaccines
; FILE REFERENCE: 469201-396
; CURRENT APPLICATION NUMBER: US/09/286,981B
; PRIOR FILING DATE: 1999-04-06
; PRIOR APPLICATION NUMBER: US 60/085,743
; PRIOR FILING DATE: 1998-05-15
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 16
; LENGTH: 414
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-286-981B-16

Query Match 55.7%; Score 1417; DB 4; Length 414;
Best Local Similarity 67.7%; Pred. No. 2e-90;
Matches 310; Conservative 46; Mismatches 46; Indels 56; Gaps 10;

Qy 38 TENEGATQVPTSSNRANESQAEQGPCKLDSERDKARKE-VVEYVKIVGESYAKSTKK 96
Db 1 TENEGSTQAATSSNNAK-----TEHRKAAKQVVEYIEKMLRE--TQDDR 44
Qy 97 RHTITVALVNLNINKNEYLNK--IVESTSESQILMMESRSKVDEAVSKFEKDSSSSS 154
Db 45 RHTQVNLNINKSAIKTYRLNVLNLEKSKDELTP---SEIKAKLDA--FEK----- 92
Qy 155 SSDSSTKPEASDTAKPNKPTPEGKVAEAEKKVVEAEKKAKDQKEEDRRNYPTITYKTLE 214
Db 93 -----EKKDTLK-----PGKVAEAEKKVVEAEKKAEQKQKEEDRRNYPTITYKTLE 138
Qy 215 LEIASDVEVKAELELVKVKANEPDEQIKQAEAEVSKAEATRLKIKITDREEAE 274
Db 139 LEIABFDVKVKEAELELVKEEAKESRNEGITIKQAEKVESKAEATRLNIKITDRKAE 198
Qy 275 EAKRADAK-----EQGPKGRAGVPGELATPDKENDAKSSDSSVGEETLPS 325
Db 199 EAKRADAKLEKANVATSDQGPGRAGVPGELATPDKENDAKSSDSSVGEETLPS 258
Qy 326 SLKPKKVAEAEKKVVEAEKKAEADQKEEDRRNYPTNTYKTLLEIAESDVEVKAELELV 385
Db 259 SLKSKKVAEAEKKVVEAEKKAEADQKEEDRRNYPTNTYKTLLEIAESDVKVEAELELV 318
Qy 386 KEEAEKPRNEEKVKQAKAEVSKAEATRLKIKITDRKAEAEAKRAAEEDVKVEKPAE 445
Db 319 KEEAEKPRDEEKIKQAKAKVESKAEATRLNENIKITDRKAEAEAKRAAEEDVKVEKPAE 378
Qy 446 QPAPAPAKAEKPAAPKPPENPAEQPKAEKPADQQAEE 483
Db 379 QPAPAPQPKAEKPAAPKPPENPAEQPKAEKPADQQAEE 414
```

```
RESULT 14
US-09-286-981B-15
; Sequence 15, Application US/09286981B
; Patent No. 6503511
; GENERAL INFORMATION:
; APPLICANT: Wizemann, Theresa M.
; APPLICANT: Koenig, Scott
; APPLICANT: Johnson, Leslie S
; TITLE OF INVENTION: Derivatives of Choline Binding Proteins for Vaccines
; FILE REFERENCE: 469201-396
; CURRENT APPLICATION NUMBER: US/09/286,981B
; PRIOR FILING DATE: 1999-04-06
; PRIOR APPLICATION NUMBER: US 60/085,743
; PRIOR FILING DATE: 1998-05-15
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 15
; LENGTH: 419
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-286-981B-15

Query Match 55.6%; Score 1416; DB 4; Length 419;
Best Local Similarity 66.7%; Pred. No. 2.3e-90;
Matches 309; Conservative 43; Mismatches 49; Indels 62; Gaps 9;

Qy 38 TENEGATQVPTSSNRANESQAEQGPCKLDSERDKARKEVEYVKIVGESYAKSTKKR 97
Db 1 TENERTQVPTSSNRGK-----PERKAAEQDFEYINKMI-----OLDKRR 41
Qy 98 HTITVALVNLNINKNEYLNKIVESTSESQILMMESRSKVDEAVSKFEKDSSSSSSD 157
Db 42 HTQNLAFNQLSRIKTEYLNGLKEK-SEAELP---SKIKAEIDAAFKQFKDT----- 90
Qy 158 SSTPEASDTAKPNKPTPEGKVAEAEKKVVEAE-----KKAQDQKEEDRRNYPTITY 210
Db 91 -----LPTPEKKVAEAEKKVVEAEKKVAEKKAKAKAEEDRRNYPTITY 136
Qy 211 KTLLEIAESDVEVKAELELVKVKANEPDEQIKQAEAEVSKAEATRLKIKITDRE 270
Db 137 KTLLEIAEFDVKVKEAELELVKKEADESRNEGITINQAKAKVESEKAEATRLKIKITDRE 196
Qy 271 EA-BEEAKRRADAKEQG---KPKGRAGVPGELATPDKENDAKSSDSSVGEETLPS 326
Db 197 KAESEAKRRADAKEQDESRRKRGKRGALGEQATPDKENDAKSSDSSVGEETLPS 256
Qy 327 LKPEKKVAEAEKKVVEAEKKAEADQKEEDRRNYPTNTYKTLLEIAESDVEVKAELELVK 386
Db 257 LKPGKVAEAEKKVVEAEADKKAQAKQKEEDRRNYPTNTYKTLLEIAESDVKVEAELELVK 316
Qy 387 EEAKEPRNEEKVKQAKAEVSKAEATRLKIKITDRKAEAEAKRAAEEDVKVEKPAEQ 446
Db 317 EEAKEPRNEEKVKQAKAEVSKAEATRLKIKITDRKAEAEAKRAAEEDVKVEKPAEQ 376
Qy 447 QPAPAPAKAEKPA-----PAPKPPENPAEQPKAEKPADQQAEE 482
Db 377 QPAPAPQPKAEKPAEPENPVPAKPPENPAEQPKAEKPADQQAEE 419

RESULT 15
US-09-286-981B-10
; Sequence 10, Application US/09286981B
; Patent No. 6503511
; GENERAL INFORMATION:
; APPLICANT: Wizemann, Theresa M.
; APPLICANT: Koenig, Scott
; APPLICANT: Johnson, Leslie S
; TITLE OF INVENTION: Derivatives of Choline Binding Proteins for Vaccines
; FILE REFERENCE: 469201-396
; CURRENT APPLICATION NUMBER: US/09/286,981B
; CURRENT FILING DATE: 1999-04-06
```

```
; PRIOR APPLICATION NUMBER: US 60/085,743
; PRIOR FILING DATE: 1998-05-15
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 414
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-286-981B-10

Query Match      55.5%; Score 1413; DB 4; Length 414;
Best Local Similarity 66.6%; Pred. No. 3.7e-90;
Matches 305; Conservative 49; Mismatches 48; Indels 56; Gaps 8;

Qy 38 TENEGATQVPTSSNRANESQAEQGEQPKLDSERDKARKE-VVEYVVKIVGESYAKSTKK 96
Db 1 TENEGSTQAATSSNMAK-----TEHRKAQKVVEYIEKMLRE--IQLDRR 44

Qy 97 RHTITVALVNLNINIKNEYLNK--IVESTSQIQLMMESRSKVDEAVSFEKDDSSSS 154
Db 45 KHTQNALNLIKLSAIKTKYLRELNVLEESKDPLP---SEIKAKLDAAFEFKIDT---- 97

Qy 155 SSDSSTKEASDTAKPNKPTPEGKVAEAKKVEAEKKAKDOKEEDRRNYPTITYKTLE 214
Db 98 -----LKPGEKVAEAKKVEAEKKAKDOKEEDRRNYPTITYKTLE 138

Qy 215 LEIAESDVEVKAELELVKVKANPRDQKIQAQAEVESQAEATRLKIKTDREBAEE 274
Db 139 LEIAEFVVKVEAELELVKVEAKESRNEGTLKQAEKVESKAEATRLNKTDRKAAEE 198

Qy 275 EAKRRADAK-----EQGPKGRAGKGVPGELATPDKKENDAKSDSSVGBETLPSP 325
Db 199 EAKRKADAKLEANVATSDQGPKGRAGKGVPGELATPDKKENDAKSDSSVGBETLPSS 258

Qy 326 SLKPEKKVAEAEKVEAEKKAEDQKEEDRRNYPTNTYKTLEIAESDVEVKAELELV 385
Db 259 SLKSGKVAEAEKVEAEKKAEDQKEEDRRNYPTNTYKTLDLEIAESDVVKVEAELELV 318

Qy 386 KEEAKEPRNEEKVKQAQAEVESKKAATRLKIKTDREKAAEEAKRKAEEEDVKKEKPAE 445
Db 319 KEEAKEPRDEEKIQAQAKVESKKAATRLNKTDRDDAEEAEAKRKAEEEDVKKEKPAE 378

Qy 446 QPQAPAPKAEKPAPKPPENPAPQPKAEKPADQQAEE 483
Db 379 QPQAPAPATQPEK--PAPKPEKPAEQPKAEKTDQQAEE 414
```

Search completed: September 28, 2005, 18:44:16
Job time : 44 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 28, 2005, 18:32:15 ; Search time 178 Seconds
(without alignments)
1475.824 Million cell updates/sec

Title: US-09-298-523D-13

Perfect score: 2546

Sequence: 1 MFASKSRKHYHSIRKFSVG.....NRLTQQPPKTEKPAQPSTP 513

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt.03.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	2546	100.0	693	2	Q97N74	Q97N74 streptococc
2	2337	99.6	693	2	Q9KK21	Q9KK21 streptococc
3	2503.5	98.3	696	2	Q9KK32	Q9KK32 streptococc
4	2489.5	97.8	516	2	Q9RQT2	Q9RQT2 streptococc
5	2420	95.1	695	2	Q9KK36	Q9KK36 streptococc
6	2070.5	81.3	684	2	Q9KK17	Q9KK17 streptococc
7	1814	71.2	695	2	Q9KK51	Q9KK51 streptococc
8	1808	71.0	681	2	Q9KK16	Q9KK16 streptococc
9	1806.5	71.0	660	2	Q9KK18	Q9KK18 streptococc
10	1803	70.8	709	2	Q9KK38	Q9KK38 streptococc
11	1796.5	70.6	487	2	Q9RQT3	Q9RQT3 streptococc
12	1780.5	69.9	752	2	Q9KK30	Q9KK30 streptococc
13	1762.5	69.2	752	2	Q9KK33	Q9KK33 streptococc
14	1756	69.0	564	2	Q69188	Q69188 streptococc
15	1756	69.0	701	2	Q9KK48	Q9KK48 streptococc
16	1752	68.8	701	2	Q8DN05	Q8DN05 streptococc
17	1752	68.8	701	2	Q9RQT5	Q9RQT5 streptococc
18	1751.5	68.8	730	2	Q9KK47	Q9KK47 streptococc
19	1744	68.5	581	2	Q33742	Q33742 streptococc
20	1741.5	68.4	680	2	Q9KK41	Q9KK41 streptococc
21	1739.5	68.3	660	2	Q9KK43	Q9KK43 streptococc
22	1735.5	68.3	730	2	Q9KK31	Q9KK31 streptococc
23	1719	67.5	699	2	Q9KK42	Q9KK42 streptococc
24	1709	67.1	655	2	Q9KK50	Q9KK50 streptococc
25	1661	65.2	657	2	Q9KK29	Q9KK29 streptococc
26	1642	64.5	684	2	Q9KK46	Q9KK46 streptococc
27	1635.5	64.2	488	2	Q9RQT6	Q9RQT6 streptococc
28	1621.5	63.7	678	2	Q9KK54	Q9KK54 streptococc
29	1619.5	63.6	678	2	Q9KK49	Q9KK49 streptococc
30	1612.5	63.3	680	2	Q9KK11	Q9KK11 streptococc
31	1608	63.2	681	2	Q9KK13	Q9KK13 streptococc

32	1599	62.8	676	2	Q9KK39	Q9KK39 streptococc
33	1599	62.8	696	2	Q9KK14	Q9KK14 streptococc
34	1581	62.1	663	2	Q30874	Q30874 streptococc
35	1418	55.7	739	2	Q9RQT4	Q9RQT4 streptococc
36	1371	53.8	820	2	Q9RQT1	Q9RQT1 streptococc
37	1363.5	53.6	869	2	Q9KK27	Q9KK27 streptococc
38	1356	53.3	929	2	Q9KK19	Q9KK19 streptococc
39	1356	53.3	929	2	Q9ZAY5	Q9ZAY5 streptococc
40	1239	48.7	565	2	Q9KK15	Q9KK15 streptococc
41	1228.5	48.3	769	2	Q9KK40	Q9KK40 streptococc
42	1215	47.7	667	2	Q9KK38	Q9KK38 streptococc
43	1195.5	47.0	770	2	Q9KK37	Q9KK37 streptococc
44	1191.5	46.8	769	2	Q9FDQ1	Q9FDQ1 streptococc
45	1189	46.7	523	2	Q33753	Q33753 streptococc

ALIGNMENTS

RESULT 1

Q97N74
ID Q97N74 PRELIMINARY; PRT; 693 AA.
AC Q97N74;
DT 01-OCT-2001 (TrEMBLrel. 18, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Choline binding protein A.
GN OrderedLocusNames=SP2190;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC BAA-334 / TIGR4;
RX MEDLINE=21357209; PubMed=11463916; DOI=10.1126/science.1061217;
RA Tettein H., Nelson K.E., Paulsen I.T., Eisen J.A., Read T.D.,
Peterson S.N., Heidelberg J.F., DeBoy R.T., Haft D.H., Dodson R.J.,
Durkin A.S., Gwinn M.L., Kolonay J.F., Nelson W.C., Peterson J.D.,
Umayam L.A., White O., Salzberg S.L., Lewis M.R., Radune D.,
RA Holtzapfel E.K., Khouri H.M., Wolf A.M., Uterback T.R., Hansen C.L.,
McDonald L.A., Feldblyum T.V., Anguoli S.V., Dickinson T.,
Hickey E.K., Holt I.E., Loftus B.J., Yang E., Smith H.O., Venter J.C.,
Dougherty B.A., Morrison D.A., Hollingshead S.K., Fraser C.M.;
"Complete genome sequence of a virulent isolate of Streptococcus pneumoniae.";
RL Science 293:498-506(2001).
DR EMBL: AE007507; AAK76241.1; -.
DR PIR: H95255; H95255.
DR HSSP: P06653; LH86.
DR TIGR: SP2190; -.
GO: GO:0016020; C:membrane; IEA.
DR InterPro: IPR002479; CW binding.
DR InterPro: IPR005877; GPOA_YSIK.
DR InterPro: IPR007756; RICH.
DR Pfam: PF01473; CW_binding_1; 8.
DR Pfam: PF05062; RICH; 1.
DR Pfam: PF04650; YSIK signal; 1.
DR TIGRFAMs: TIGR01168; YSIK_signal; 1.
KW Complete proteome.
SQ SEQUENCE 693 AA; 77762 MW; 6F8F47C32E344A41 CRC64;

Query Match 100.0%; Score 2546; DB 2; Length 693;

Best Local Similarity 100.0%; Pred. No. 2.3e-80;

Matches 513; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MFASKSRKHYHSIRKFSVGVSVMVSLVNGSVVHATENEGATQVPTSSNRANESQAEQ 60

Db 1 MFASKSRKHYHSIRKFSVGVSVMVSLVNGSVVHATENEGATQVPTSSNRANESQAEQ 60

Qy 61 GEQPKLDSERDKARKEVEEVKKIVGESYAKSKKRHTITVALVNEUNNKIYLNKIV 120

Db 61 GEQPKLDSERDKARKEVEEVKKIVGESYAKSKKRHTITVALVNEUNNKIYLNKIV 120


```

Db 61 GQPKLDSERDKARKEVEYVKIVGESYAKSTKQRHTITVALVNLNINIKNEYNKIV 120
QY 121 ESTSESQILMMESRSKVDEAVSKFEKDSOSSSSSSSTKPEASDTAKPNKPTPEGKV 180
Db 121 ESTSESQILMMESRSKVDEAVSKFEKDSOSSSSSSSTKPEASDTAKPNKPTPEGKV 180
QY 181 ABAKKVBEAEAKKADQKEEDRRNPTTYKTLELEIAESDVVEVKAELELVKVKANEP 240
Db 181 ABAKKVBEAEAKKADQKEEDRRNPTTYKTLELEIAESDVVEVKAELELVKVKANEP 240
QY 241 DEOKIKQAEAVESKQAEATRLKIKTDREAEAEAKRADAKEOGKP---KGRKRGV 296
Db 241 DEOKIKQAEAVESKQAEATRLKIKTDREAEAEAKRADAKEOGKPPKRRKGRKRGV 300
QY 297 PGELATPDKKENDAKSSSSSVGEETLPSPSLKPEKKVAEAKKVEAEKKAEDQKEEDRR 356
Db 301 PGEQATPDKKENDAKSSSSSVGEETLPSPSLKPEKKVAEAKKVEAEKKAEDQKEEDRR 360
QY 357 NYPTTYKTLELEIAESDVVEVKAELELVKESAKPRNEEKVKQAEVSKAEATRL 416
Db 361 NYPTTYKTLELEIAESDVVEVKAELELVKESAKPRNEEKVKQAEVSKAEATRL 420
QY 417 KIKTDKKAEEAKKAEDVKKEKPAEQOPAPAPKAEKPAPAPKPNPAEQPIAEKP 476
Db 421 KIKTDKKAEEAKKAEDVKKEKPAEQOPAPAPKAEKPAPAPKPNPAEQPIAEKP 480
QY 477 ADQQAEDYARSEBEYNRLTQQQPPKTEKPAQSTP 513
Db 481 -DQQAEDYARSEBEYNRLTQQQPPKTEKPAQSTP 516

```

RESULT 4

```

Q9ROT2 ID Q9ROT2 PRELIMINARY; PRT; 516 AA.
AC Q9ROT2;
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Hypothetical protein pspC (Fragment).
GN Name=pspC;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=L81905.
RX MEDLINE=2003819; PubMed=10569772;
RA Brooks-Walter A., Briles D.E., Hollingshead S.K.;
RT "The pspC gene of Streptococcus pneumoniae encodes a polymorphic
RT protein, PspC, which elicits cross-reactive antibodies to PspA and
RT provides immunity to pneumococcal bacteremia.";
RL Infect. Immun. 67, 6533-6542 (1999).
DR EMBL; AF068649; AAF13459.1; -.
DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR005877; Gpos_Ysirk.
DR InterPro; IPR007756; RICH.
DR Pfam; PF05062; RICH; 1.
DR Pfam; PF04650; Ysirk_signal; 1.
DR TIGRFAMs; TIGR01168; Ysirk_signal; 1.
KW Hypothetical protein.
FT NON_TER 516
SQ SEQUENCE 516 AA; 57886 MW; 1119782688BB7E35 CRC64;

```

```

Query Match 97.8%; Score 2489.5; DB 2; Length 516;
Best Local Similarity 98.2%; Pred. No. 1.6e-78;
Matches 504; Conservative 2; Mismatches 6; Indels 1; Gaps 1;

QY 1 MFASKSERKVHYSIRKFSVGVSVVVASLVMSGVVHATENEGATQVPTSSNRANESQAEQ 60
Db 1 MFASKSERKVHYSIRKFSVGVSVVVASLVMSGVVHATENEGLPST-TSSNRANESQAEQ 59

```

```

QY 61 GEOPKLLDSERDKARKEVEYVKIVGESYAKSTKQRHTITVALVNLNINIKNEYNKIV 120
Db 60 GEOPKLLDSERDKARKEVEYVKIVGESYAKSTKQRHTITVALVNLNINIKNEYNKIV 119
QY 121 ESTSESQILMMESRSKVDEAVSKFEKDSOSSSSSSSTKPEASDTAKPNKPTPEGKV 180
Db 120 ESTSESQILMMESRSKVDEAVSKFEKDSOSSSSSSSTKPEASDTAKPNKPTPEGKV 179
QY 181 ABAKKVBEAEAKKADQKEEDRRNPTTYKTLELEIAESDVVEVKAELELVKVKANEP 240
Db 180 ABAKKVBEAEAKKADQKEEDRRNPTTYKTLELEIAESDVVEVKAELELVKVKANEP 239
QY 241 DEOKIKQAEAVESKQAEATRLKIKTDREAEAEAKRADAKEOGKPKGRKRGVPGEL 300
Db 240 DEOKIKQAEAVESKQAEATRLKIKTDREAEAEAKRADAKEOGKPKGRKRGVPGEL 299
QY 301 ATPDKENDAKSSSSSVGEETLPSPSLKPEKKVAEAKKVEAEKKAEDQKEEDRRNPT 360
Db 300 ATPDKENDAKSSSSSVGEETLPSPSLKPEKKVAEAKKVEAEKKAEDQKEEDRRNPT 359
QY 361 NYTKTLELEIAESDVVEVKAELELVKESAKPRNEEKVKQAEVSKAEATRL 420
Db 360 NYTKTLELEIAESDVVEVKAELELVKESAKPRNEEKVKQAEVSKAEATRL 419
QY 421 DRKAAEEAKKAEDVKKEKPAEQOPAPAPKAEKPAPAPKPNPAEQPIAEKP 480
Db 420 DRKAAEEAKKAEDVKKEKPAEQOPAPAPKAEKPAPAPKPNPAEQPIAEKP 479
QY 481 AEDYARSEBEYNRLTQQQPPKTEKPAQSTP 513
Db 480 AEEYARSEBEYNRLTQQQPPKTEKPAQSTP 512

```

RESULT 5

```

Q9KK36 ID Q9KK36 PRELIMINARY; PRT; 695 AA.
AC Q9KK36;
DT 01-OCT-2000 (TRENBLrel. 15, Created)
DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
DE Surface protein PspC.
GN Name=pspC;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=G394;
RX MEDLINE=21888621; PubMed=11891047; DOI=10.1016/S0378-1119(01)00896-4;
RA Iannelli F., Oggioni M.R., Pozzi G.;
RT "Allelic variation in the highly polymorphic locus pspC of
RT Streptococcus pneumoniae.";
RL Gene 284, 63-71 (2002).
DR EMBL; AF154023; AAF73792.1; -.
DR HSP; P06653; IH8G.
DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR002479; CW_binding.
DR InterPro; IPR005877; Gpos_Ysirk.
DR InterPro; IPR007756; RICH.
DR Pfam; PF01473; CW_binding_1; 8.
DR Pfam; PF05062; RICH; 1.
DR Pfam; PF04650; Ysirk_signal; 1.
DR TIGRFAMs; TIGR01168; Ysirk_signal; 1.
SQ SEQUENCE 695 AA; 78361 MW; F8C84F08C4EFFFBA CRC64;

```

```

Query Match 95.1%; Score 2420; DB 2; Length 695;
Best Local Similarity 95.6%; Pred. No. 4.9e-76;
Matches 494; Conservative 6; Mismatches 11; Indels 6; Gaps 2;

QY 1 MFASKSERKVHYSIRKFSVGVSVVVASLVMSGVVHATENEGATQVPTSSNRANESQAEQ 60
Db 1 MFASKSERKVHYSIRKFSVGVSVVVASLVMSGVVHATENEGATQVPTSSNRANESQAEQ 60

```

```
QY 61 GEQPKLDSERDKARKEVEEYVKIIVGESYAKSTKGRHTITVALVNLNINIKNEYLNKIV 120
DB 61 GEQPKLDSERDKARKEVEEYVKIIVGESYAKSTKGRHTITVALVNLNINIKNEYLNKIV 120
QY 121 ESTSESQIILMMESRSKVDVAVSKFEKDSSSSSSSSSSTKPEASDTAKPNKPTPEGEKV 180
DB 121 ESTSESQIILMMESRSKVDVAVSKFEKDSSSSSSSSSSTKPEASDTAKPNKPTPEGEKV 180
QY 181 AEAKKVEEAEKKAQKQEDRRNYPTIYTKTLEIAESDVEVKKAELVLVKYKANEPR 240
DB 181 AEAKKVEEAEKKAQKQEDRRNYPTIYTKTLEIAESDVEVKKAELVLVKYKANEPR 240
QY 241 DEQIKQAEAEVESKQAEATRLKKITDREAEEAEKRRDAKEQG----KPKGRAKRGV 296
DB 241 DEQIKQAEAEVESKQAEATRLKKITDREAEEAEKRRDAKEQGESSKPKRSVRKGD 300
QY 297 PGEIATPDKKENDAKSSDSSVGEETLPSLSLKPEKKVAEAKKVEEAEKKAEDQKEDRR 356
DB 301 LGEQATPDKKENDAKSSDSSVGEETLPSLSLKPEKKVAEAKKVEEAEKKAEDQKEDRR 360
QY 357 NYPNTYTKTLEIAESDVEVKKAELVLKVEEAEKPRNEEKVKQAKAEVSKKAEATRL 416
DB 361 NYPNTYTKTLEIAESDVEVKKAELVLKVEEAEKPRNEEKVKQAKAEVSKKAEATRL 420
QY 417 KIKTRDKKAEAEKKAEDKVKKEPAEQQPAPAPAKPAKPENPAEQPKAEKP 476
DB 421 KIKTRDKKAEAEKKAEDKVKKEPAEQQPAPAPAKPAKPENPAEQPKAEKP 478
QY 477 ADQQAEDYARRSEENRLTQQQPPKTEKPAQPSPT 513
DB 479 ADQQAEDYARRSEENRLTQQQPPKTEKPAQPSPT 515
```

RESULT 6

```
Q9KKL17 PRELIMINARY; PRT; 684 AA.
AC Q9KKL17;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Surface protein PspC.
GN Name=pspC;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=srf22;
RX MEDLINE=21888621; PubMed=11891047; DOI=10.1016/S0378-1119(01)00896-4;
RA Iannelli F., Oggioni M.R., Pozzi G.;
RT "Allelic variation in the highly polymorphic locus pspC of
RL Streptococcus pneumoniae.";
RL Gene 284:63-71(2002).
RL EMBL; AF154039; AAF73811.1; -.
DR HSSP; P06653; 1GMV.
DR GO; GO:0016020; C.membrane; IEA.
DR InterPro; IPR002479; CW_binding.
DR InterPro; IPR005877; Gpos_Ysirk.
DR InterPro; IPR007756; RICH.
DR Pfam; PF01473; CW_binding_1; 7.
DR Pfam; PF05062; RICH; 1.
DR Pfam; PF04650; Ysirk_signal; 1.
DR TIGRFAMs; TIGR01168; Ysirk_signal; 1.
SQ SEQUENCE 684 AA; 77589 MW; 20FB892F82FDDF25 CRC64;
```

Query Match 81.3%; Score 2070.5; DB 2; Length 684;
Best Local Similarity 81.0%; Pred. No. 4.8e-64;

Matches 427; Conservative 34; Mismatches 49; Indels 17; Gaps 4;

QY 1 MFASKSERKHYHSIRKFSIGVASVAVSLVMGVSVHHATENEGTQTPTSSNRANESQAEQ 60

```
DB 1 MFASKSERKHYHSIRKFSIGVASVAVSLVMGVSVHHATENEGTQTPTSSNRANESQAEQ 58
QY 61 GEQPKLDSERDKARKEVEEYVKIIVGESYAKSTKGRHTITVALVNLNINIKNEYLNKIV 120
DB 59 -EQRLELDLDRDKVKEVREYKVKVLYSKTSRHKKTVDIVNKLQINNEYLNKII 117
QY 121 ESTSE-SQIILMMESRSKVDVAVSKFEKDSSSSSSSSSSTKPEASDTAKPNKPTPEGEK 179
DB 118 QSTSTYEELQKLMMESQSEVDKAVSEFKDLSSSSSSSSSTPEASDTAKPNKPTPELEKK 177
QY 180 VAERKKVVEAEKKAQKQEDRRNYPTIYTKTLEIAESDVEVKKAELVLVKYKANEPR 239
DB 178 VAERKKVVEAEKKAQKQEDRRNYPTIYTKTLEIAESDVEVKKAELVLVKYKANEPR 237
QY 240 RDEQIKQAEAEVESKQAEATRLKKITDREAEEAEK----RRADAEQOGPKGRAK 293
DB 238 RDEQIKQAEAEVESKQAEATRLKKITDREAEEAEKLEAVEKNAATSEOGPKPKRVK 297
QY 294 RGVPEIATPDKKENDAKSSDSSVGEETLPSLSLKPEKKVAEAKKVEEAEKKAEDQKEE 353
DB 298 RGALGEQATPDKKENDAKSSDSSVGEETLPSLSLKPEKKVAEAKKVEEAEKKAEDQKEE 357
QY 354 DRNYPNTYTKTLEIAESDVEVKKAELVLKVEEAEKPRNEEKVKQAKAEVSKKAEAT 413
DB 358 DRNYPNTYTKTLEIAESDVEVKKAELVLKVEEAEKPRNEEKVKQAKAEVSKKAEAT 417
QY 414 RLEKIKTRDKKAEAEKKAEDKVKKEPAEQQPAPAPAKPAKPENPAEQPKAEKP 466
DB 418 RLEKIKTRDKKAEAEKKAEDKVKKEPAEQQPAPAPAKPAKPENPAEQPKAEKP 477
QY 467 PAEQPKAEKPADQQAEDYARRSEENRLTQQQPPKTEKPAQPSPT 513
DB 478 PAEQPKAEKPADQQAEDYARRSEENRLTQQQPPKTEKPAQPSPT 524
```

RESULT 7

```
Q9KKK51 PRELIMINARY; PRT; 695 AA.
AC Q9KKK51;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Surface protein PspC.
GN Name=pspC;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC6307;
RX MEDLINE=21888621; PubMed=11891047; DOI=10.1016/S0378-1119(01)00896-4;
RA Iannelli F., Oggioni M.R., Pozzi G.;
RT "Allelic variation in the highly polymorphic locus pspC of
RL Streptococcus pneumoniae.";
RL Gene 284:63-71(2002).
RL EMBL; AF154009; AAF73776.1; -.
DR HSSP; P06653; 1HCX.
DR GO; GO:0016020; C.membrane; IEA.
DR InterPro; IPR002479; CW_binding.
DR InterPro; IPR005877; Gpos_Ysirk.
DR InterPro; IPR007756; RICH.
DR InterPro; IPR000533; Tropomyosin.
DR Pfam; PF01473; CW_binding_1; 10.
DR Pfam; PF05062; RICH; 1.
DR Pfam; PF04650; Ysirk_signal; 1.
DR PRINTS; PR00194; TROPOMYOSIN.
DR TIGRFAMs; TIGR01168; Ysirk_signal; 1.
SQ SEQUENCE 695 AA; 78521 MW; F208BC7285DA4FCE CRC64;
```

Query Match 71.2%; Score 1814; DB 2; Length 695;

Best Local Similarity 73.8%; Pred. No. 3.1e-55;

Matches 381; Conservative 40; Mismatches 51; Indels 44; Gaps 5;

QY 1 MFASKSERKVHYSIRKFSVGVASVAVSLVMGVSVVHATENEGATQVPTSSNRANESQAEQ 60
 DB 1 MFASKSERKVHYSIRKFSVGVASVAVSLVMGVSVVHATENEGSTQAATSSNNAK----- 54
 QY 61 GEQPKKLDSEDRKARKEVEEVVKITVGSYAKSTKKRHTITVALVNLNINKVEYLNKIV 120
 DB 55 -----TEHWKAQVDEYIKKKL-----QLDRKHTQNVGLTKLVKTEYHLGLS 101
 QY 121 ESTSESQILMMESRSKVDEAVSKFEKSDSSSSSSSTKPEASDTAKPNKPTPEGSKV 180
 DB 102 VSKSEAE-LPSEIKAKLDAAFQPKKOT-----LPTPEGKKV 139
 QY 181 AEAKKVEAEAKKADQKEEDRRNYPTTYKTLELEIAESDVEVKAELELVKVKANEP 240
 DB 140 AEAEKKVEAEAKKADQKEKDLRNYPTTYKTLELDIAESDVEVKAELELVKKEAKESR 199
 QY 241 DEQIKQAEEVESKQAEATRLKKTDRREAEAEAKRADAKQEG---KPKGRKRGVP 297
 DB 200 DEKKINQAKVSEQAATRLKKTDRKAEAEAKRADAKQEDSKRRKSRVKGDL 259
 QY 298 GELATPDKKENDAKSSDSVGBETLPSLSLKPEKVAEAEKKEVBAEKKAEDQKEEDRRN 357
 DB 260 GEPATPDKKENDAKSSDSVGBETLPSLSLKPGKVAEAEKKEVBAEKKAEDQKEEDHRN 319
 QY 358 YPTNTYKTLELEIAESDVEVKAELELVKKEAEKPRNEEKVKQAKAEVESKKAETRLK 417
 DB 320 YPTNTYKTLELEIAESDVEVKAELELVKKEAKGSRNEEKVKQAKAEVESKKAETRLK 379
 QY 418 IKTDKPKAEAEAKRAAEEDVKKEPAQOPAPAPAPAPAPAPAPAPAPAPAPAPAPAP 477
 DB 380 IKTDKPKAEAEAKRAAEEDVKKEPAQOPAPAPAPAPAPAPAPAPAPAPAPAPAPAP 439
 QY 478 DQAEEDYARSEEEYNRLTQQPPKTEKPAQPS 513
 DB 440 DQAEEDYARSEEEYNRLTQQPPKTEKPAQPS 475

RESULT 8

Q9KK16 PRELIMINARY; PRT; 681 AA.
 AC Q9KK16;
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
 DE Surface protein PspC.
 GN Name=pspC;
 OS Streptococcus pneumoniae.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1313;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=serf25;
 RX MEDLINE=21888621; PubMed=11891047; DOI=10.1016/S0378-1119(01)00896-4;
 RA Iannelli F., Oggioni M.R., Pozzi G.;
 RT "Allelic variation in the highly polymorphic locus pspC of
 Streptococcus pneumoniae.";
 RL Gene 284:63-71(2002).
 DR EMBL; AF154040; AAF73812.1; -.
 DR HSSP; P06653; 1HXC.
 DR GO; GO:0016020; C:membrane; IEA.
 DR InterPro; IPR002479; CW binding.
 DR InterPro; IPR005877; Gpce YSIK.
 DR InterPro; IPR007756; RICH-
 DR Pfam; PF01473; CW binding_1; 8.
 DR Pfam; PF05062; RICH; 1.
 DR Pfam; PF04650; YSIK signal; 1.
 DR TIGRFAMs; TIGR01168; YSIK signal; 1.
 SQ SEQUENCE 681 AA; 76728 MW; 38FE4782653D51A8 CRC64;

Query Match 71.0%; Score 1808; DB 2; Length 681;
 Best Local Similarity 70.8%; Pred. No. 5e-55;

Matches 380; Conservative 44; Mismatches 51; Indels 62; Gaps 7;
 QY 1 MFASKSERKVHYSIRKFSVGVASVAVSLVMGVSVVHATENEGATQVPTSSNRANESQAEQ 60
 DB 1 MFASKSERKVHYSIRKFSVGVASVAVSLVMGVSVVHATEKEVTTQVPTSSNRANKSQ--- 57
 QY 61 GEQPKKLDSEDRKARKEVEEVVKITVGSYAKSTKKRHTITVALVNLNINKVEYLNKIV 120
 DB 58 -----TEHWKAQVDEYIKKKL-----QLDRKHTQNVGLTKLVKTEYHLGLS 104
 QY 121 ESTSESQILMMESRSKVDEAVSKFEKSDSSSSSSSTKPEASDTAKPNKPTPEGSKV 180
 DB 105 VSKSEAE-LPSEIKAKLDAAFQPKKOT-----LPTPEGKKV 142
 QY 181 AEAKKVEAEAKKADQKEEDRRNYPTTYKTLELEIAESDVEVKAELELVKVKANEP 240
 DB 143 AEAEKKVEAEAKKADQKEKDLRNYPTTYKTLELDIAESDVEVKAELELVKGGYKPR 202
 QY 241 DEQIKQAEEVESKQAEATRLKKTDRREAEAEAKRADAK-----EQGKP 288
 DB 203 DEQVNOAKVSEQAATRLKKTDRREAEAEAKRADAK-----EQGKP 262
 QY 289 KGRKRGVPGELATPDKKENDAKSSDSVGBETLPSLSLKPEKVAEAEKKEVBAEKKA 348
 DB 263 KRGKRGALGEOATPDKKENDAKSSDSVGBETLPSLSLKPEKVAEAEKKEVBAEKKA 322
 QY 349 DQKEEDRRNYPTNTYKTLELEIAESDVEVKAELELVKKEAEKPRNEEKVKQAKAEVESK 408
 DB 323 DQKEEDRRNYPTNTYKTLELEIAESDVEVKAELELVKKEAEKPRNEEKVKQAKAEVESK 382
 QY 409 KAEATRLKIKTDKPKAEAEAKRAAEEDVKKEPAQOPAPAPAPAPAPAPAPAPAPAP 461
 DB 383 KAEATRLKIKTDKPKAEAEAKRAAEEDVKKEPAQOPAPAPAPAPAPAPAPAPAPAP 442
 QY 462 PKPENPAQPKAEKPADQAEEDYARSEEEYNRLTQQPPKTEKPA-----QPSTP 513
 DB 443 PKPENPAQPKAEKPADQAEEDYARSEEEYNRLTQQPPKTEKPA-----QPSTP 499

RESULT 9

Q9KK18 PRELIMINARY; PRT; 660 AA.
 AC Q9KK18;
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
 DE Surface protein PspC.
 GN Name=pspC;
 OS Streptococcus pneumoniae.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1313;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=serf2;
 RX MEDLINE=21888621; PubMed=11891047; DOI=10.1016/S0378-1119(01)00896-4;
 RA Iannelli F., Oggioni M.R., Pozzi G.;
 RT "Allelic variation in the highly polymorphic locus pspC of
 Streptococcus pneumoniae.";
 RL Gene 284:63-71(2002).
 DR EMBL; AF154038; AAF73810.1; -.
 DR HSSP; P06653; 1H8G.
 DR GO; GO:0016020; C:membrane; IEA.
 DR InterPro; IPR002479; CW binding.
 DR InterPro; IPR005877; Gpce YSIK.
 DR InterPro; IPR007756; RICH-
 DR Pfam; PF01473; CW binding_1; 8.
 DR Pfam; PF05062; RICH; 1.
 DR Pfam; PF04650; YSIK signal; 1.
 DR TIGRFAMs; TIGR01168; YSIK signal; 1.
 SQ SEQUENCE 660 AA; 74515 MW; 374EB582FD09A659 CRC64;

Query Match 71.0%; Score 1806.5; DB 2; Length 660;

Best Local Similarity 73.6%; Pred. No. 5.5e-55;
Matches 384; Conservative 39; Mismatches 48; Indels 51; Gaps 7;
Qy 1 MFASKSERKVVHSIRKFSVGVASVVVSLVMSVGHATENEGATQVPTSSNRANESQAEQ 60
Db 1 MFASKSERKVVHSIRKFSVGVASVVVSLVMSVGHATENEGATQVPTSSNRANESQAEQ 57
Qy 61 GEOPKKLDSDERDKARKEVEYVKIVGESYAKSTKKRHTITVALVNLNNIKN 120
Db 58 -----TEYKAAKQVDEYITKKL-----QLDRREHIQNVGLLTKLVKTEYLHRLS 104
Qy 121 ESTSESQIILMMESRSKVDVAESKPKDSSSSSSSSSTKPEASDRAKNKPTPEGKIV 180
Db 105 VSKEKSEAE-LPSIKAKLDAAFELFKDT-----LPTPEGKKV 142
Qy 181 AEAKKVEEAEKKAADQKEEDRRNYPTTYKTLELEIAESDVEYKKAELVVKYKANEPR 240
Db 143 AEAKKVEEAEKKAADQKEEDRRNYPTTYKTLELEIAESDVEYKKAELVVKYKANEPR 202
Qy 241 DEQIKQAEAEVESKQAEATRLKKIKTDREAEAEAKRRADAK-----EQGPKGR 291
Db 203 DEKINQAKAVENKKAETELKNIKTDREKA-EAEKRRADAKLQEANVATSEODKSKRR 261
Qy 292 AKRGVPELATPDKKENDAKSSDSVSGEETLPSPSLKPEKVAEAEKVKKAESDQK 351
Db 262 ANREVGLATPDKKENDAKSSDSVSGEETLPSPSLKPEKVAEAEKVKKAESDQK 321
Qy 352 EEDRRNYPTTYKTLELEIAESDVEYKKAELVVKYKANEPRNEEKVKQAKAEVESKKA 411
Db 322 EEDRRNYPTTYKTLELEIAESDVEYKKAELVVKYKANEPRNEEKVKQAKAEVESKKA 381
Qy 412 ATRLEKIKTDRAKAEAEKKAABEDKVKKEPAEQPAPAPKAEKPAKPNPAEQ 471
Db 382 ATRLEKIKTDRAKAEAEKKAABEDKVKKEPAEQPAPAPKAEKPAKPNPAEQ 438
Qy 472 KAEKPAEQABEDYARRSEEEYNRLTQQPPKTEKPAQPSPTP 513
Db 439 KAEKPAEQABEDYARRSEEEYNRLTQQPPKTEKPAQPSPTP 480

RESULT 10
Q9KK38 PRELIMINARY; PRT; 709 AA.
AC Q9KK38
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Surface protein PspC.
GN Name=pspC;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=g387;
RX MEDLINE=21888621; PubMed=11891047; DOI=10.1016/S0378-1119(01)00896-4;
RA Iannelli F., Oggioni M.R., Pozzi G.;
RT "Allelic variation in the highly polymorphic locus pspC of
Streptococcus pneumoniae";
RL Gene 284:63-71(2002).
DR EMBL; AF154021; AAF73790.1; -.
DR HSSP; P06653; 1H8G.
DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR002479; CW_binding.
DR InterPro; IPR005877; Gpos_Ysirk.
DR InterPro; IPR007756; RICH.
DR InterPro; IPR000533; Tropomyosin.
DR Pfam; PF01473; Tropomyosin.
DR Pfam; PF05062; RICH; 1.
DR Pfam; PF04650; Ysirk_signal; 1.
DR PRINTS; PR00194; TROPOMYSIN.
DR TIGRFAMs; TIGR01168; Ysirk_signal; 1.

SQ SEQUENCE 709 AA; 80251 MW; 32BBC9E380EBB7A CRC64;
Query Match 70.8%; Score 1803; DB 2; Length 709;
Best Local Similarity 71.1%; Pred. No. 7.6e-55;
Matches 382; Conservative 49; Mismatches 54; Indels 52; Gaps 9;
Qy 1 MFASKSERKVVHSIRKFSVGVASVVVSLVMSVGHAT-ENEGATQVPTSSNRANESQ-- 57
Db 1 MFASKSERKVVHSIRKFSVGVASVVVSLVMSVGHAT-ENEGATQVPTSSNRANESQTE 60
Qy 58 -----AEOGEPKKLDSDERDKARKEVEYVKIVGESYAKSTKKRHTITVALVNLNNIKN 113
Db 61 HMAAEGQKHRRIDLRNKAIDYEIEKMLSE--IQDKRKHQNFALNLSRIKT 118
Qy 114 EYLN--KIVESTSESOLQI---LMMESRSKVDVAESKPKDSSSSSSSSSTKPEASDTA 168
Db 119 EYLYGLSVLKESEAEELPSKAEELPSKIKELTAAPFEHFKDT----- 160
Qy 169 KPNKPTPEGKVAEAEKKAADQKEEDRRNYPTTYKTLELEIAESDVEYKKA 228
Db 161 -----LRPGEKVAEAEKKAADQKEEDRRNYPTTYKTLELEIAESDVEYKKA 215
Qy 229 LELVKYKANEPRNEOKIKQAEAEVESKQAEATRLKKIKTDREAEAEAKRRADAK----- 283
Db 216 LELVKYKANEPRNEOKIKQAEAEVESKQAEATRLKKIKTDREAEAEAKRRADAK----- 275
Qy 284 -----EQGPKGRAGVPELATPDKKENDAKSSDSVSGEETLPSPSLKPEKVAE 336
Db 276 EKMTATSEQKPKGRAGVPELATPDKKENDAKSSDSVSGEETLPSPSLKPEKVAE 335
Qy 337 EKKVEAEKKAADQKEEDRRNYPTTYKTLELEIAESDVEYKKAELVVKYKANEPRNE 396
Db 336 EKKVEAEKKAADQKEEDRRNYPTTYKTLELEIAESDVEYKKAELVVKYKANEPRNE 395
Qy 397 KVKQAKAEVESKQAEATRLKKIKTDREAEAEAKRRADAK-----EAKRKA 456
Db 396 KVKQAKAEVESKQAEATRLKKIKTDREAEAEAKRRADAK-----EAKRKA 454
Qy 457 KPAKPAKPNPAEQABEDYARRSEEEYNRLTQQPPKTEKPAQPSPTP 513
Db 455 K--PAPKPEKPAEHPKAEENPADQQAEDYARRSEEEYNRLTQQPPKTEKPAQPSPTP 509

RESULT 11
Q9RQT3 PRELIMINARY; PRT; 487 AA.
AC Q9RQT3
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein pspC (Fragment).
GN Name=pspC;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BG8090;
RX MEDLINE=20038319; PubMed=10569772;
RA Brooks-Walter A., Briles D.E., Hollingshead S.K.;
RT "The pspC gene of Streptococcus pneumoniae encodes a polymorphic
protein, pspC, which elicits cross-reactive antibodies to PspA and
provides immunity to pneumococcal bacteremia";
RL Infect. Immun. 67:6533-6542(1999).
DR EMBL; AF068648; AAF13458.1; -.
DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR005877; Gpos_Ysirk.
DR InterPro; IPR007756; RICH.
DR Pfam; PF05062; RICH; 1.
DR Pfam; PF04650; Ysirk_signal; 1.
DR TIGRFAMs; TIGR01168; Ysirk_signal; 1.
KW Hypothetical protein.

```
FT NON_TER 487 487
SQ SEQUENCE 487 AA; 54962 MW; 7BDBB21C65341E86 CRC64;

Query Match
Best Local Similarity 72.1%; Score 1796.5; DB 2; Length 487;
Matches 379; Conservative 44; Mismatches 50; Indels 53; Gaps 6;

QY 1 MFASKSERKHYISIRKFSVGVASVVASLVMSGVVHATENEGATQVPTSSNRANESQAEQ 60
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1 MFASKNERKHYISIRKFSIGVASVAVSLFMSGVVHATEKVTTQVATSFNKANKSQ--- 57

QY 61 GQPKKLDSEDKARKEVEYVKIVGVSYAKSTKKRHTITVALNVLNNIKNEVLN 120
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
58 -----TEHWKAQVDEYITKKL-----QLDRKHTQNVGLTKLGVIKTEYLHRLS 104

QY 121 ESTSESQILMMESRSKVDEAVSKFEDSSSSSSSDSTKPEASDTAKPNKPTPEGKV 180
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
105 VSKEKSEAE-LPSEIKAKLDAAEQPKD-----LPTPEGKV 142

QY 181 ABAKKVEAEAKKADQKEEDRRNPTTYKTLEIAESDVVEVKAELELVKVKANEP 240
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
143 ABAAKVEAEAKKADQKEEDRRNPTTYKTLEIAESDVVEVKAELELVKKEAGSR 202

QY 241 DSQIKQAEAVESKQAEATRLKKIKTDREAE-----BEAKERADAEQ 286
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
203 NEQVNAQAKAVESKQAEATRLKKIKTDREAEATRLKIKTDREAEAKKADAKSQD 262

QY 287 KPGKRAKGVPGELATPDKKENDAKSSDSSVGEETLPSPSLKPEKKVAEAEKK 346
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
263 ESKRVKGVGVEQATLDKENDAKSSDSSVGEETLPSPSLKSGKKVAEAEKKVAEAK 322

QY 347 AEDQKEEDRRNPTTYKTLEIAESDVVEVKAELELVKKEAKPRNEEKVKQAEVE 406
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
323 AKDQKEEDRRNPTTYKTLEIAESDVVEVKAELELVKKEAKSRNEEKVKQAEVE 382

QY 407 SKKAEATRLKIKTDREAEAKKAAEEDKVEKPAEQPOPAKPAKPAKPN 466
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
383 SKKAEATRLKIKTDREAEAKKAAEEDKVEKPAEQPOPAKPAKPAKPN 441

QY 467 PAEQKPAKPADQQAEDYARRSEEEYNRLTQQQPPKTEKPAQST 512
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
442 PAEQKPAKPADQQAEDYARRSEEEYNRLTQQQPPKTEKPAQST 487
```

RESULT 12

```
O9KK30
ID O9KK30 PRELIMINARY; PRT; 752 AA.
AC O9KK30;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Surface protein PspC.
GN Name=pspC;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]_TaxID=1313;
RP SEQUENCE FROM N.A.
RC STRAIN=9408;
RX MEDLINE=21888621; PubMed=11891047; DOI=10.1016/S0378-1119(01)00896-4;
RA Iannelli F., Oggioni M.R., Pozzi G.;
RT "Allelic variation in the highly polymorphic locus pspC of
RT Streptococcus pneumoniae.";
RL Gene 284:63-71(2002).
DR EMBL; AF154029; AAF73798.1; -.
DR HSSP; P06653; 1H8G.
DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR002479; CW binding.
DR InterPro; IPR005877; GpOs_YsIRK.
DR InterPro; IPR007756; RICH_YsIRK.
DR Pfam; PF01473; CW binding_1; 12.
DR Pfam; PF05062; RICH; 1.
```

RESULT 13

```
O9KK33
ID O9KK33 PRELIMINARY; PRT; 752 AA.
AC O9KK33;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Surface protein PspC.
GN Name=pspC;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]_TaxID=1313;
RP SEQUENCE FROM N.A.
RC STRAIN=940;
RX MEDLINE=21888621; PubMed=11891047; DOI=10.1016/S0378-1119(01)00896-4;
RA Iannelli F., Oggioni M.R., Pozzi G.;
RT "Allelic variation in the highly polymorphic locus pspC of
RT Streptococcus pneumoniae.";
RL Gene 284:63-71(2002).
DR EMBL; AF154026; AAF73795.1; -.
DR HSSP; P06653; 1HCX.
DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR002479; CW binding.
DR InterPro; IPR005877; GpOs_YsIRK.
DR InterPro; IPR007756; RICH_YsIRK.
DR Pfam; PF01473; CW binding_1; 13.
DR Pfam; PF01473; CW binding_1; 13.
```



```
DR InterPro; IPR005877; Gpos.YSIRK.
DR InterPro; IPR007756; RICH.
DR Pfam; PF01473; CW_binding_1; 10.
DR Pfam; PF05062; RICH; 1.
DR Pfam; PF04650; YSIRK signal; 1.
DR TIGRFAMs; TIGR01168; YSIRK signal; 1.
SQ SEQUENCE 701 AA; 79098 MW; A80E31FE2846F1BF CRC64;

Query Match      69.0%; Score 1756; DB 2; Length 701;
Best Local Similarity 70.9%; Pred. No. 3.le-53;
Matches 372; Conservative 50; Mismatches 47; Indels 56; Gaps 8;

Qy 1 MFASKSERKRVHSYIRKFSVGVASVVVSLVVGSVVHATENEGATQVPTSSNRANESQAEQ 60
Db 1 MFASKSERKRVHSYIRKFSVGVASVVVSLVVGSVVHATENEGSTQATSSNNAK----- 54

Qy 61 GEQPKKLDSEDRDKARKE-VEEYVKKIVGESYAKSTKKRHTITVALVNLNINIKNEYLNK- 118
Db 55 -----TEHRKAQVVDYIEKMLRE--IQDDRKHQTQNALNINIKLSAIKTKYLREL 104

Qy 119 -IVSTSESQILMMESRSKVDEAVSKPEKSDSSSSSSSTKPEASDTAKPNKPTPG 177
Db 105 NVLBEKSKDELFP---SEIKAKLDAAFEXFKDT-----LKPQ 138

Qy 178 EKVAEAKKKVVEAEKKAQKQKEEDRRNYPTITYKTLELEIAESDVEVKKAELVLKVKAN 237
Db 139 EKVAEAKKKVBEAKKKAQKQKEEDRRNYPTITYKTLELEIAEFQVVKKEALELVKEAK 198

Qy 238 EPRDEQIKQAEEVESQAEATRLKIKTDREAEAEAKRRADAK-----EOGKP 288
Db 199 ESRNEGTTIKQAEKVESKKAATRLNIKTDKKAEEAEAKKADAKLKEANVATSDQGP 258

Qy 289 KGRAKGVPGELATPDKKENDAKSDSSVGBETLPSPLKPEKKVAEAEKKVVEAEKKA 348
Db 259 KGRAKGVPGELATPDKKENDAKSDSSVGBETLPSSSLKSGKKVAEAEKKVVEAEKKA 318

Qy 349 DQKEEDRRNYPTITYKTLELEIAESDVEVKKAELVLKKEAEKPRNEKVKQAEVESK 408
Db 319 DQKEEDRRNYPTITYKTLDLEIAESDVVKGAELVLKKEAEKPRDEEKIKQAKAKVESK 378

Qy 409 KAEATRLKIKTDKKAEEAEAKKAAEEDKVKEPAEQOPAPAPAKPAKPENPA 468
Db 379 KAEATRLNIKTDKKAEEAEAKKAAEEDKVKEPAEQOPAPAPATQPEK--PAKP 436

Qy 469 EQPKAEKPADQQAEDYARRSEEEYNRLTQOOPPKTEKPAQFSTP 513
Db 437 EQPKAEKTDQQAEDYARRSEEEYNRLTQOOPPKTEKPAQFSTP 481
```

Search completed: September 28, 2005, 18:42:39
Job time : 181 secs

THIS PAGE IS BLANK

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 28, 2005, 18:33:03 ; Search time 42 Seconds
(without alignments)
1175.219 Million cell updates/sec

Title: US-09-298-523D-13
Perfect score: 2546
Sequence: 1 MFASKSERKVVHYSIRKFSVG.....NRLTQQQPKTEKPAQPSTP 513

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 79:*
1: Pirl:*
2: pirl2:*
3: pirl3:*
4: pirl4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2546	100.0	693	2 H95255	choline binding pr
2	1756	69.0	701	2 H98120	choline binding pr
3	496	19.5	619	2 A97887	surface protein ps
4	496	19.5	619	2 A41971	surface protein ps
5	397	15.6	744	2 F95013	pneumococcal surfa
6	393.5	15.5	1110	2 I51116	NF-180 - sea lampr
7	365.5	14.4	1164	1 FCSOAG	IGA Fc receptor pr
8	365	14.3	1020	1 QFHUH	neurofilament trip
9	361.5	14.2	1701	2 T09127	probable erythrocy
10	360.5	14.2	1134	2 A60234	IGA Fc receptor pr
11	343.5	13.5	606	2 A43427	neurofilament trip
12	338	13.3	1052	1 A44937	kinetoplast-associ
13	336.5	13.2	3488	2 T34418	hypothetical prote
14	334.5	13.1	6642	2 T29757	protein UNC-89 - C
15	329.5	12.9	1390	2 S51364	sperm tail-specifi
16	321	12.6	539	2 A28549	M24 protein precur
17	318.5	12.5	522	2 C96608	hypothetical prote
18	317	12.5	706	2 A45990	functional sarcopl
19	314.5	12.4	1072	1 A37221	neurofilament trip
20	310.5	12.2	854	2 S02003	neurofilament trip
21	310	12.2	699	2 E84565	hypothetical prote
22	309.5	12.2	5327	2 T13564	microtubule-associ
23	300.5	11.8	501	2 A44643	M protein precursor
24	299.5	11.8	771	1 A33430	h-caldesmon - chic
25	297	11.7	1087	1 QFMNH	neurofilament trip
26	297	11.7	1815	2 C81169	IGA-specific metal
27	296.5	11.6	1094	2 S49313	protein kinase - s
28	294	11.5	858	2 S15762	neurofilament trip
29	292.5	11.5	1526	2 A45605	mature-parasite-in

30	291.5	11.4	729	2 S68191	triadin - human
31	290.5	11.4	1128	2 T30296	R27-2 protein - Tr
32	290	11.4	1773	2 A81937	IGA-specific metal
33	287	11.3	849	2 S00030	neurofilament trip
34	283	11.1	1192	2 A71623	probable secreted
35	281.5	11.1	644	2 S55395	neurofilament prot
36	281.5	11.1	916	2 A27864	neurofilament trip
37	280.5	11.0	587	2 JC1419	Fc gamma (IgG) rec
38	279.5	11.0	532	2 S54871	M protein - Strept
39	276.5	10.9	667	2 A40713	cylicin I - bovine
40	276	10.8	488	2 T46014	cylicin II - bovin
41	274.5	10.8	564	2 A60115	M protein precursor
42	273.5	10.7	2464	1 QRMSPI	microtubule-associ
43	273	10.7	332	2 B43427	neurofilament prot
44	270.5	10.6	598	2 B40713	cylicin I - human
45	270	10.6	483	2 A26297	M6 protein - Strep

ALIGNMENTS

RESULT 1

H95255
choline binding protein A [imported] - Streptococcus pneumoniae (strain TIGR4)
C:Species: Streptococcus pneumoniae
C>Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 09-Jul-2004
C:Accession: H95255

R:Tettelein, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Heide
on, J.D.; Umayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapple, I.
nson, T.; Hickey, E.K.; Holt, I.E.
Science 293, 498-506, 2001

A:Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison,
A:Title: Complete Genome Sequence of a Virulent isolate of Streptococcus pneumoniae.
A:Reference number: A95000; MUID:21357209; PMID:11463916

A:Accession: H95255

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-693 <KUR>

A:Cross-references: UNIPROT:Q9N74; GB:AE005672; PIDN:AAK76241.1; PID:g14973701; GSPDB:G1

A:Experimental source: strain TIGR4

C:Genetics:

A:Gene: SP2190

Query Match 100.0%; Score 2546; DB 2; Length 693;
Best Local Similarity 100.0%; Pred. No. 3.3e-89;
Matches 513; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MFASKSERKVVHYSIRKFSVG	VASV	VVVASLV	MGSVV	VHATENEGATQVPTSSNRANESQAEQ	60
Db	1	MFASKSERKVVHYSIRKFSVG	VASV	VVVASLV	MGSVV	VHATENEGATQVPTSSNRANESQAEQ	60
Qy	61	GEQPKLDSERDKARKEVEEYVK	IVGESYAKSTKGRHTITVALV	NLNINIKNEYL	NKIV	120	
Db	61	GEQPKLDSERDKARKEVEEYVK	IVGESYAKSTKGRHTITVALV	NLNINIKNEYL	NKIV	120	
Qy	121	ESTSESQILIMESRSKVD	EA	VSKFEKDS	SSSSSSSSSTKPEASD	TAKPNKPTPEGSKV	180
Db	121	ESTSESQILIMESRSKVD	EA	VSKFEKDS	SSSSSSSSSTKPEASD	TAKPNKPTPEGSKV	180
Qy	181	AEAKKVEEAEKKA	DKOKEEDRRNYPITYKTLE	IAESDVEVK	KAELV	KVKANPR	240
Db	181	AEAKKVEEAEKKA	DKOKEEDRRNYPITYKTLE	IAESDVEVK	KAELV	KVKANPR	240
Qy	241	DEQIKQAEAEVSKQAEAT	RLKKIKTDREAEAEK	RRADAK	EQGKPKGR	KRGVPGEL	300
Db	241	DEQIKQAEAEVSKQAEAT	RLKKIKTDREAEAEK	RRADAK	EQGKPKGR	KRGVPGEL	300
Qy	301	ATPDKENDAKSS	SSSGVEETLPSFSLKPEK	VAAEKV	EA	KKKAAEDQKEEDRRNYP	360
Db	301	ATPDKENDAKSS	SSSGVEETLPSFSLKPEK	VAAEKV	EA	KKKAAEDQKEEDRRNYP	360
Qy	361	NTYKTLLEIAESD	VEVKKAELV	KBEAKESPRNEEK	VKQAKAEV	SKKAEATLEKTKT	420
Db	361	NTYKTLLEIAESD	VEVKKAELV	KBEAKESPRNEEK	VKQAKAEV	SKKAEATLEKTKT	420

Db 361 NTYKTLLEIAESDVVVKKAELELVKEEAKSPRNEEKVKQAKAEVSKAEATRLKLEIKT 420
Qy 421 DRKKAEEBAKKAEEEDKVKPEAPQOPAPAPKAEKAPAPKAPENPAEQPKAEKPADOQ 480
Db 421 DRKKAEEBAKKAEEEDKVKPEAPQOPAPAPKAEKAPAPKAPENPAEQPKAEKPADOQ 480
Qy 481 AEEYARRSEYENRLTQQQPPKTEKPAQPSSTP 513
Db 481 AEEYARRSEYENRLTQQQPPKTEKPAQPSSTP 513

RESULT 2
H98120
choline binding protein A [imported] - Streptococcus pneumoniae (strain R6)
C;Species: Streptococcus pneumoniae
C;Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 09-Jul-2004
C;Accession: H98120
R;Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; DeHoff, B.S.; E
e, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.; M
y, P.; Sun, P.M.; Winkler, M.E.
J. Bacteriol. 183, 5709-5717, 2001
A;Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.R.;
A;Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.
A;Reference number: A97872; MUID:21429245; PMID:11544234
A;Accession: H98120
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-701 <KUR>
A;Cross-references: UNIPROT:Q9KK48; UNIPROT:Q8DN05; GB:AE007317; PID:AA00797.1; PID:gl
C;Genetics:
A;Gene: pspC

Query Match 69.0%; Score 1756; DB 2; Length 701;
Best Local Similarity 70.9%; Pred. No. 1.6e-59;
Matches 372; Conservative 50; Mismatches 47; Indels 56; Gaps 8;

Qy 1 MFASKSERKHYISIRKFSVGVASVAVSLVMSVGHATENEGATQVPTSSNRANESQAEQ 60
Db 1 MFASKSERKHYISIRKFSVGVASVAVSLVMSVGHATENEGSTOATSSNMAK----- 54
Qy 61 GEQPKLDSERDKARKE-VEYVVKIVGESVAKTKKRRHTITVALVNLNNIKNEYLNK- 118
Db 55 -----TEHRKAQKVVDYIEIKMRE--IQDRLRRKHTONVALNTKLSAIKTKYLREL 104
Qy 119 -IVESTSESQILMMESRSKVDKAVSKFEDSSSSSSSSSTKPEASDTAKPNKPTBPG 177
Db 105 NVLEKSKDELPA---SEIKALDAAFEKFKDPT-----LKPQ 138
Qy 178 EKVAEAKKKVEEAKKADQKEEDRRNYPTITYKTLELEIAESDVVKKAELELVKVKAN 237
Db 139 EKVAEAKKKVEEAKKAEQKEEDRRNYPTNTYKTLELEIAEFQVVKAELELVKEEAK 198
Qy 238 EPREQIKQAEAEVSKQAEATRLKKITDREAEBAEKARRADAK-----EQKP 288
Db 199 ESRNEGITQAKAEVSKAEATLENIKTRDKAEBAEKAKADAKLKEANVATSDQGP 258
Qy 289 KGRAKGVPGELATPDKKENDAKSSDSVGBETLPSPLKPEKVKAEKVKAEKKA 348
Db 259 KGRAKGVPGELATPDKKENDAKSSDSVGBETLPSSSLKGGKVKAEKVKAEKKA 318
Qy 349 DQKEEDRRNYPTNTYKTLELEIAESDVVKKAELELVKEEAKPRNEEKVKQAKAEVSK 408
Db 319 DQKEEDRRNYPTNTYKTLDLSIESDVVKVKEAELELVKEEAKPRDEKIKQAKAEVSK 378
Qy 409 KAEATRLKIKITDRKKAEEBAKKAEEEDKVKPEAPQOPAPAPKAEKAPAPKAPENPA 468
Db 379 KAEATRLKIKITDRKKAEEBAKKAEEEDKVKPEAPQOPAPAPQPEK--PAPKPEKPA 436
Qy 469 EQPKAEKPADQAEEDYARRSEYENRLTQQQPPKTEKPAQPSSTP 513
Db 437 EQPKAEKTDQQAEDYARRSEYENRLTQQQPPKTEKPAQPSSTP 481

RESULT 3
A97887
surface protein pepA precursor [imported] - Streptococcus pneumoniae (strain R6)
C;Species: Streptococcus pneumoniae
C;Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 09-Jul-2004
C;Accession: A97887
R;Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; DeHoff, B.S.; E
e, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.; M
y, P.; Sun, P.M.; Winkler, M.E.
J. Bacteriol. 183, 5709-5717, 2001
A;Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.R.;
A;Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.
A;Reference number: A97872; MUID:21429245; PMID:11544234
A;Accession: A97887
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-619 <KUR>
A;Cross-references: UNIPROT:Q54972; UNIPROT:Q8DR10; GB:AE007317; PID:AAK98925.1; PID:gl
C;Genetics:
A;Gene: pspA

Query Match 19.5%; Score 496; DB 2; Length 619;
Best Local Similarity 33.7%; Pred. No. 3.2e-12;
Matches 168; Conservative 62; Mismatches 148; Indels 120; Gaps 21;

Qy 26 VASLVMSGVVHATENEGATQVPTSSNRANESQAEQEQPKLDSERDKARKEVEEYVKI 85
Db 14 VALLGAGFVASQPTVVRAEESPV-----ASQSKAEKDYDAAKQDAK--NAKKAVEDAQKAL 67
Qy 86 VGSYAKSTKKRHTITVALVNE---LNNIKNEYLNKIVESTSESQILMMESRSKVD 142
Db 68 ---DDAKAAQKYDEDOQKTEEKAALFEKAAAESEMDKAVAAVQQAYL-----AYQQA 115
Qy 143 VSKFEKDSSSSSSSSTKPEAS-----DTAKPNKPTBPGKVAEAKKKVEEAKKADQK 198
Db 116 TDKAADAADAKMIDEAKKREBEAKTKFNTVRAMVVPEP-SQLAETKKSEAKQKA----- 170
Qy 199 EEDRRNYPTITYKTLELEIAESDVVKKAELELVKVKANEPREDEQIKQAEAEVESKQAE 258
Db 171 -----PELT-KKLEBAKAEKAEKATEAKQKQVDAEVAPOAKIAELENQVHRLEQE 222
Qy 259 ATRIKKITDREAEBAEKARRADAKGEQKPKGRAKRGVPGELATPDKKENDAKSSDSV 318
Db 223 ---LKEI--DESESEDYAKGFRAPLOSK-----LDAAKAKLS-- 255
Qy 319 EETLPSPLKPEKVKAEAKKVEEAKKADQKEEDRRNYPTNTYKTLELEIAESDVVK 378
Db 256 -----KLEELSKIDE-----LDAEIAKLELDQK 279
Qy 379 KAELELVKEEAKSPRN-EKVKQKA-KAEVSKKAEATRLKIKITDRKKAEEBAEKKA-AE 435
Db 280 AAE-----ENNVEDYFKGLEKTIKAAKAE--LEKTEADLKAVNEPEKPA 327
Qy 436 EDKVKPEAPQOPAPAPKAEKAPAPKAPENPAEQPKAEKPADQQAEDYARRSEYENR 495
Db 328 ETPAPEAPAEQPKPAAPAP---QAPAPKPEKPAEQPKPEKTDDQQAEDYARRSEYENR 384
Qy 496 LTQOQPPKTEKPA-QPST 512
Db 385 LTQOQPPKAEKPAAPAKPT 402

RESULT 4
A41971
surface protein pepA precursor - Streptococcus pneumoniae
N;Alternate names: pneumococcal surface protein A
C;Species: Streptococcus pneumoniae
C;Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
C;Accession: A41971; A60282; A33134
R;Yother, J.; Briles, D.E.
J. Bacteriol. 174, 601-609, 1992
A;Title: Structural properties and evolutionary relationships of PepA, a surface protein
A;Reference number: A41971; MUID:92105030; PMID:1729249

C:Superfamily: neurofilament triplet H protein
C:Keywords: coiled coil; heterotrimer; intermediate filament; nerve; phosphoprotein
F:1-100/Domain: amino-terminal <NTD>
F:101-410/Domain: rod #status predicted <ROD>
F:411-1020/Domain: carboxyl-terminal <CTD>
F:502-826/Region: 14-residue repeats
F:503,511,518,526,532,540,546,552,560,566,574,580,586,594,600,606,614,620,628,634,640,646
(covalent) #status predicted
F:732,768/Binding site: phosphate (Thr) (covalent) #status predicted

Query Match 14.3%; Score 365; DB 1; Length 1020;
Best Local Similarity 25.6%; Pred. No. 4.2e-07;
Matches 150; Conservative 77; Mismatches 247; Indels 112; Gaps 20;

QY 5 KSERKVH--SIRKPSVGASVVVASLVGMSVYVHATENEGATQVPTSSNRANESQAEQGE 62
DB 439 KSEIKIVKVESEKTVIVEEQTEQTQVTEVEEKEKEBEQGEGBEGEEBAQGE 498
QY 63 QPKLDSERDKA--RKEVEEYK-KIVGESYAKSTKKRHTITVALVNLNLIKNEYLNKI 119
DB 499 EETKSPPAEAAASPEKEAKSPVKEEAKSPAEEAKSPEKEAKSPAEEVKSPEKAKSP----- 553
QY 120 VESTSESQILQMLMESRSKVDKAVSKFEKDSSTSSSSSSSTKPEASDTAKPNKPT-EPGE 178
DB 554 AKEEAKSPPEAKSPKEEAKSPAEEVKSPEKAKSPAEEAKSPAEEAKSPAEEAKSPAEEAKSPA 613
QY 179 KVAEAKKVVEAKKAKDOKEEDRRNYPTITYKTLEIAESDVEVKKAELELVKVEAKPRNEEKV--Q 398
DB 614 SPAEAKSPVKEEAKSPAEEVKSPEKAKSP-----KKEAKSPAEEAKSPAEEAKSPAEEAKSPA 662
QY 239 PRDEQIKIQAEEVSK-----QAATRLIKITD-REAEBAEAKRRADAKQOGPKGRA 292
DB 663 PEKAKSPVKEEAKSPAEEVKSPEKAKSPAEEVKSPEKAKSPAEEVKSPEKAKSPAEEVKSPEKA 722
QY 293 KRGVPGELATPDKNDAKSSDSSVGEETLPSPLKPEKVAEAKKVEAKKAEADQKE 352
DB 723 KSPVKEAKTPKAKSPAEEVKSPEKAKSPAEEVKSPEKAKSPAEEVKSPEKAKSPAEEVKSPEKA 775
QY 353 EDRRNYPTITYKTLEIAESD-----VEVKAELELVKVEAKPRNEEKV-----KV 398
DB 776 SPADKPEKAKSPAEEVKSPEKAKSPAEEVKSPEKAKSPAEEVKSPEKAKSPAEEVKSPEKAKSPA 835
QY 399 KQAKAEVSKKAEATRLKIKITDRKAEBAEAKKAEADQKVEK----- 442
DB 836 KEPPKAEBAEAPATPKTEKKDSKK--EAPKKEAPKPVKEEKEAPKPVKEEKEAPKPVKEEKEAPK 893
QY 443 --PAEQPAPAPAKAEKP-----APAKPENPAEQPAPAEK 476
DB 894 KEEADKKKVPPEKAEAPKAEVKEADKPKTEVAKPEDDAKAEKSPKAEKKEA-AP 952
QY 477 ADQQAEEYARRSEBYNRLTQQQPPKTEKPA-----OPSTP 513
DB 953 EKKDTKEEKAKKPEK-----PKTEAKAKEDDKTLSEKPSKP 989

RESULT 9
T09127
probable erythrocyte-binding protein MAEBL - Plasmodium yoelii
C:Species: Plasmodium yoelii
C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 09-Jul-2004
C:Accession: T09127
R:Kapke, S.H.I.; Noe, A.R.; Fraser, T.S.; Blair, P.L.; Adams, J.H.
Proc. Natl. Acad. Sci. U.S.A. 95, 1230-1235, 1998
A:Title: A family of chimeric erythrocyte binding proteins of malaria parasites.
A:Reference number: Z16577; MUID:98115903; PMID:9448314
A:Accession: T09127
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1701 <KAP>
A:Cross-references: UNIPROT:O61164; EMBL:AF031886; NID:g2947227; PID:g2947228
A:Experimental source: subspecies yoelii; strain YM
C:Genetics:
A:Gene: maeb1

A:Introns: 62/1; 1648/1; 1674/2; 1697/1
C:Keywords: alternative splicing; cell binding; erythrocyte invasion

Query Match 14.2%; Score 361.5; DB 2; Length 1701;
Best Local Similarity 29.6%; Pred. No. 9e-07;
Matches 142; Conservative 75; Mismatches 168; Indels 95; Gaps 19;

QY 57 QAEQGEQPKLDSERDKAR--KEVEEYKIVGSIYAKSTKKRHTITVALVNLNLI--- 111
DB 1107 EAQAAWAKAEAEERKKAEEVKAEEERKRIEAEKAEERKRIEAEKAEERKRIEAE 1166
QY 112 -KNEVLNKIVESTESQILQMLMESRSKVDKAVSKFEKDSSTSSSSSSSTKPEASDTAKP 170
DB 1167 KAEERKIIIEAAKAE-----EERKRIEAE-----KAEERKIIIEAA----- 1205
QY 171 NKPTPEGEKVAAEKVKVEAEKKA--DQKEEDRRNYPTITYKTLE-----LEIAESDVEVK 225
DB 1206 -KAEERKKAEEVKAEEAKKAEAEKKAEEKKAEEAKKAEAEKKAEEAKKAEAEKKAEEAKK 1264
QY 226 KAELELVKVAEPRDEQIKIQA--EAEVESKQAEATRLKIKITDRKAEBAEAKRRADAK 283
DB 1265 K-----KAEAKKAEAEKKAEEAKKAEAEKKAEEAKKAEAEKKAEEAKKAEAEKKAEEAKK 1306
QY 284 EQGPKGPAKEGVPCELATPDKNDAKSSDSSVGEETLPSPLKPEKVAEAKKVEEA 343
DB 1307 LQRPKPKGKRSKESRRRKGRKLGQRKQ-----KQKRLKQRQKQKK 1353
QY 344 KQAEADQKEEDRRNYPTITYKTLEIAESDVEVKKAELELVKVEAKPRNEEKV--Q 400
DB 1354 KKRLKQRKQRKKEKKAEEAKKAEAEKRIEAE-KAEERKKAEEAKKAEAEKRIEAE 1412
QY 401 KAEVESKKAATRLKIKITDRKAEBAEAKR-----KAEEDK-----VKEPAEQPAP 451
DB 1413 KAEAEERKRIEAE-----EKAEAEERKRIEAEKAEERKRIEAEVKAEEERKRIEAE 1463
QY 452 APKAEKPAKPENPAEQPAPAEKPAEQQAEE-----DYARRSEBYNRLTQQQPPKTEK 506
DB 1464 EKAEERKRIEAEVKAEEERKRIEAEKAEERKRIEAEKAEAEKRIEAEKAEAEKRIEAE 1521

RESULT 10
A60234
Iga Fc receptor precursor - Streptococcus agalactiae (strain SB35)
N:Alternate names: Iga-binding protein; protein Bac
N:Contains: beta antigen
C:Species: Streptococcus agalactiae
C:Date: 08-Dec-1992 #sequence_revision 08-Dec-1992 #text_change 09-Jul-2004
C:Accession: A60234; S14595; A60230
R:Heden, L.O.; Frithz, E.; Lindahl, G.
Eur. J. Immunol. 21, 1481-1490, 1991
A:Title: Molecular characterization of an Iga receptor from group B streptococci: sequen
ents with Iga-binding capacity.
A:Reference number: A60234; MUID:91257158; PMID:2044657
A:Accession: A60234
A:Molecule type: DNA
A:Residues: 1-1134 <HED>
A:Cross-references: UNIPROT:Q99051; EMBL:X58470; NID:g46520; PIDN:CRA41384.1; PID:g46521
A:Note: the source is designated as group B streptococcus strain SB35
R:Heden, L.; Frithz, E.; Lindahl, G.
Submitted to the EMBL Data Library, March 1991
A:Description: Molecular characterization of an Iga receptor from group B streptococci: s
fragments.
A:Reference number: S14595
A:Accession: S14595
A:Molecule type: DNA
A:Residues: 1-1134 <HE2>
A:Cross-references: EMBL:X58470; NID:g46520; PIDN:CRA41384.1; PID:g46521
A:Note: the source is designated as Streptococcus agalactiae
R:Lindahl, G.; Akerstrom, B.; Vaerman, J.P.; Stenberg, L.
Eur. J. Immunol. 20, 2241-2247, 1990
A:Title: Characterization of an Iga receptor from group B streptococci: specificity for
A:Reference number: A60230; MUID:91055597; PMID:2242758
A:Accession: A60230

QY 429 AK-----RKAABEDVK--EKPAQOPAPAPAKPAKPAKPENPAEQPKA 473
DB 1823 LKSPTKESKSPSTTKTGDSKESKSPKPKPK-SPTPK-KSPGSPK-KKSKSPEA 1879
QY 474 EKPADQQAED 484
DB 1880 EKPPAPKLT RD 1890

RESULT 15
S51364
sperm tail-specific protein mst101(2) - fruit fly (Drosophila hydei)
C:Species: Drosophila hydei
C:Date: 19-Jul-1996 #sequence_revision 26-Jul-1996 #text_change 09-Jul-2004
C:Accession: S51364; S34154
R:Neesen, J.; Padmanabhan, S.; Buenemann, H.
Eur. J. Biochem. 225, 1089-1095, 1994
A:Title: Tandemly arranged repeats of a novel highly charged 16-amino-acid motif represe
lpha-helical rods within the extremely elongated spermatozoa of Drosophila hydei.
A:Reference number: S51364; MUID:95045538; PMID:7957199
A:Accession: S51364
A:Status: nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1-1390 <NEE>
A:Cross-references: UNIPROT:Q08696; EMBL:X73481
R:Neesen, J.; Heinlein, U.A.O.; Buenemann, H.
submitted to the EMBL Data Library, June 1993
A:Reference number: S34153
A:Accession: S34154
A:Molecule type: DNA
A:Residues: 1-163, 'E', 164-236, 'Q', 237-254, 257-320, 'E', 321-1390 <NEW>
A:Cross-references: EMBL:X73481; NID:G313201; PID:G313202
C:Genetics:
A:Gene: mst101(2)
A:Cross-references: FlyBase:FBgn0011816

Query Match 12.9%; Score 329.5; DB 2; Length 1390;
Best Local Similarity 26.5%; Pred. No. 1.2e-05;
Matches 137; Conservative 81; Mismatches 224; Indels 75; Gaps 14;

QY 39 ENEGATQVPTSSNRANE-SQAEQGEQPKLDSER---DKARKEVEYVKKIVGESYAKST 94
DB 362 EKKACKELAKKKKGADEKKECEEAANKKKAERKKCEKKAERKEAAEKKCEAAKKE 421
QY 95 KKRHTITVALVNLNNIKNEYLNKIVBSTSBSQILMMESRSKVDKAVSKFEDSSSS 154
DB 422 KE-----AAERKCEELAKNKKAAEK--KCKEAAKKEKAAERKKCEELAKK 468
QY 155 SSDSSTKPEASDTAKPNKPTPEGKVAFAKKKVBEAE-----KKAKDQKE----- 199
DB 469 IKKAAEKKKCBETAKGKGEVAERKKCEELAKKIKKABIKKKCKKLAKKEREKTAEEKKCEK 528
QY 200 EDRNYPITITVKTLELETAESDVEVKAELELVKANEPRDEQIKQA-----EAEVES 254
DB 529 AAKKKKGAAEKKKCEKAAKKKCEKKAERKKCEKSAKKKKAERKKCEKKAERKEAAEK 588
QY 255 KQABATRLKIKITDREAEAEERADAKQCKPKGRAK-----RGVPGELATPDCKE 307
DB 589 KCEEAARKEVEAERKKCEELAKKIKKAAEKKKCKEAAKKEKAAERKKCEELAKKIKKA 648
QY 308 NDAKSSSSSVGEETLPSPSLKPE-----KKVAEAKKVBEA--KKKAEDQKEE 353
DB 649 AEKKKKCKLAKKETAEKKKCEKAAKKKCEKKAERKKCEKKAERKKCEKKAERKEAAEK 708
QY 354 DRRNYPNTYKTLLELETAESDVEVKAELELVKAEKEPRNEEKVQAKAEVESKKAERAT 413
DB 709 EKEAAERKKCEELAKKTK-----KAAEKKKCKLAKKKKAGKGNKLGKNGKGA--- 759
QY 414 RLEKIKTRKKAEEAKKAAEEDKVKEPQAPAPAPKAPKAPAPKAPKAPKAPKAPKAPK 473
DB 760 -----LKEKKCKRELAKKGAEEKKCEKKAERKKCEKKAERKKCEKKAERKEA 812
QY 474 EKPADQQAEDYARRSEE-EYNRLTQQQPPKTEKPAQ 509

DB 813 EKTAKKKKAAEKKKKCKEAAKKGKKEAEKKKCEKTAK 849
Search completed: September 28, 2005, 18:43:32
Job time : 49 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 28, 2005, 18:29:20 ; Search time 171 Seconds
(without alignment)

1160.281 Million cell updates/sec

Title: US-09-298-523D-13

Perfect score: 2546

Sequence: 1 MFASKSRKHVHSRKPSVG.....NRLTQQPPKTEKPAQSTP 513

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_l6Dec04:.*
1: geneseqpl1980s:.*
2: geneseqpl1990s:.*
3: geneseqpl2000s:.*
4: geneseqpl2001s:.*
5: geneseqpl2002s:.*
6: geneseqpl2003as:.*
7: geneseqpl2003bs:.*
8: geneseqpl2004s:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2546	100.0	693	5	AAU76890 Pneumococ
2	2546	100.0	693	6	ABU02720 S. pneumo
3	2546	100.0	693	8	ADM92271 S. pneumon
4	2546	100.0	694	3	AAy81653 Streptoco
5	2366	92.9	655	2	AAy32099 Choline b
6	2366	92.9	655	2	AAy49226 CbpA of s
7	2366	92.9	655	7	ABW01575 Streptoco
8	2366	92.9	655	8	ADRI10795 Choline b
9	2268	89.1	460	2	AAy49250 N-termina
10	2268	89.1	460	8	ADRI10831 Choline b
11	2265	89.0	460	2	AAy32189 N-termina
12	2251	88.4	459	2	AAy32190 N-termina
13	2251	88.4	459	2	AAy49251 N-termina
14	2251	88.4	459	8	ADRI10832 Choline b
15	2210	86.8	446	2	AAy49143 Amino aci
16	2193	86.1	446	2	AAy49140 Amino aci
17	2117	83.2	428	2	AAy32110 Choline b
18	2117	83.2	428	2	AAy49238 Choline b
19	2117	83.2	428	7	ABW01587 Streptoco
20	2117	83.2	428	8	ADRI10817 Choline b
21	1995	78.4	406	2	AAy32098 Choline b
22	1995	78.4	406	2	AAy49225 Polypepti
23	1995	78.4	406	7	ABW01574 Streptoco
24	1995	78.4	406	8	ADRI10794 Choline b
25	1796.5	70.6	487	5	AAU76889 Pneumococ

26	1756	69.0	564	2	AAW62654	Aaw62654	C3 bindin
27	1756	69.0	701	6	ABU46304	Abu46304	Protein e
28	1752	68.8	701	5	AAU76888	Aau76888	Pneumoco
29	1744	68.5	581	2	AAy43394	Aay43394	S. pneumo
30	1653	64.9	437	8	ADR10833	Adr10833	Choline b
31	1635.5	64.2	488	5	AAU76887	Aau76887	Pneumoco
32	1581	62.1	663	2	AAy32104	Aay32104	Choline b
33	1581	62.1	663	2	AAy43395	Aay43395	S. pneumo
34	1581	62.1	663	2	AAy49232	Aay49232	CbpA of s
35	1581	62.1	663	5	AAE29201	Aae29201	Streptoco
36	1581	62.1	663	5	ABG60639	Abg60639	Partial s
37	1581	62.1	663	6	ABP55319	Abp55319	Streptoco
38	1581	62.1	663	7	ABW01581	Abw01581	Streptoco
39	1581	62.1	663	7	ADL199571	Adl199571	GST-pigr
40	1581	62.1	663	8	ADRI10801	Adri10801	Choline b
41	1580	62.1	419	2	AAy32182	Aay32182	N-termina
42	1580	62.1	419	2	AAy49243	Aay49243	N-termina
43	1580	62.1	419	8	ADRI10824	Adri10824	Choline b
44	1524	59.9	406	2	AAy49152	Aay49152	Amino aci
45	1517.5	59.6	631	2	AAW46444	Aaw46444	CbpA, a c

ALIGNMENTS

RESULT 1
AAU76890
ID AAU76890 standard; protein; 693 AA.
XX
AC AAU76890;
XX
DT 05-JUN-2002 (first entry)
XX
DE Pneumococcal surface protein C19 TIGR (PspC19 TIGR).
XX
KW Hic; factor H; fH; antibacterial; PspC19 TIGR;
KW pneumococcal surface protein C19 TIGR.
XX
OS Streptococcus pneumoniae.
XX
FN WO200208426-A2.
XX
PD 31-JAN-2002.
XX
PF 19-JUL-2001; 2001WO-EP008409.
XX
PR 20-JUL-2000; 2000SE-00002738.
XX
PA (HANS-) HANSA MEDICAL AB.
XX
PI Bjorck L, Sjöholm A, Janulczyk R, Pozzi G, Iannelli F;
XX
DR WPI; 2002-257337/30.
XX
PT Polypeptide having ability to bind factor H, and proteins and peptides
PT derived from polypeptides useful as vaccine for infections caused by
PT bacteria and for identifying agents inhibiting binding of factor H to
protein.
XX
PS Disclosure; Page 59-61; 63pp; English.
XX
CC The invention relates to the Streptococcus pneumoniae Hic polypeptide
CC having the ability to bind factor H (fH). The Hic protein and the
CC polynucleotide encoding it can be used to make vaccine compositions
CC capable of generating an immune response to Streptococcus pneumoniae or
CC binding to an anti-protein Hic antibody. The protein and its homologues
CC are useful for identifying an agent which inhibits binding of factor H to
CC Streptococcus pneumoniae comprising incubating any one of the
CC the polypeptide with factor H and a test agent, monitoring binding of fH to
CC the polypeptide and determining thereby whether the test agent inhibits
CC binding of factor to the polypeptide. This sequence represents the
CC pneumococcal surface protein C19 TIGR (PspC19 TIGR), a factor H binding
CC protein


```
XX SQ Sequence 693 AA;
Query Match 100.0%; Score 2546; DB 5; Length 693;
Best Local Similarity 100.0%; Pred. No. 1.8e-152;
Matches 513; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MFASKSERKVHYSIRKFSVGASVVVASLVMSGVVHATENEGATQVPTSSNRANESQAEQ 60
Db 1 MFASKSERKVHYSIRKFSVGASVVVASLVMSGVVHATENEGATQVPTSSNRANESQAEQ 60

Qy 61 GEOPKLDSDRKARKEVEYVKIVGESYAKSTKKRHTITVALVNLNINKNEYLNKIV 120
Db 61 GEOPKLDSDRKARKEVEYVKIVGESYAKSTKKRHTITVALVNLNINKNEYLNKIV 120

Qy 121 ESTSESQQLILMMESRSKVDPAVSKFPEKDSOSSSSSSDSTKPEASDTAKPNKPTPEGEKV 180
Db 121 ESTSESQQLILMMESRSKVDPAVSKFPEKDSOSSSSSSDSTKPEASDTAKPNKPTPEGEKV 180

Qy 181 AEAKKVVEAEKKAQKQKEEDRRNYPITYTKTLEIAESDVEVKKAELVLVKVANEPR 240
Db 181 AEAKKVVEAEKKAQKQKEEDRRNYPITYTKTLEIAESDVEVKKAELVLVKVANEPR 240

Qy 241 DEQIKIQAEAVESKQAEATRLKKIKTDREAEAEAKRRADAKSQGPKGRKRGVPGEL 300
Db 241 DEQIKIQAEAVESKQAEATRLKKIKTDREAEAEAKRRADAKSQGPKGRKRGVPGEL 300

Qy 301 ATPDKKENDAKSSDSSVGEETLPSPLKPEKKVAEAEKKVVEAKKKAEDQKEEDRRNYP 360
Db 301 ATPDKKENDAKSSDSSVGEETLPSPLKPEKKVAEAEKKVVEAKKKAEDQKEEDRRNYP 360

Qy 361 NTYKTLLEIAESDVEVKKAELVLKVEAKPEPRNEEKVKQAKAVESKKAETRLKIKT 420
Db 361 NTYKTLLEIAESDVEVKKAELVLKVEAKPEPRNEEKVKQAKAVESKKAETRLKIKT 420

Qy 421 DRKXAEAEAKKAAEDDKVKEKPAEQPPAPAPKAEKPAPAPKPPENPAEQPKAEKPADQQ 480
Db 421 DRKXAEAEAKKAAEDDKVKEKPAEQPPAPAPKAEKPAPAPKPPENPAEQPKAEKPADQQ 480

Qy 481 AEEDYARRSEEEYNRLTQQQPPKTEKPAQPS 513
Db 481 AEEDYARRSEEEYNRLTQQQPPKTEKPAQPS 513
```

RESULT 2
ABU02720
ID ABU02720 standard; protein; 693 AA.

```
XX AC ABU02720;
XX DT 23-OCT-2003 (revised)
XX DT 11-FEB-2003 (first entry)
XX DE S. pneumoniae type 4 strain protein from coding region #2299.
XX KW Bacterial meningitis; pneumonia; sepsis; otitis media; ear infection;
XX KW antiinflammatory; antibacterial; immunostimulant; auditory; respiratory;
XX KW gene therapy; vaccine.
XX OS Streptococcus pneumoniae; type 4 strain.
XX PN WO200277021-A2.
XX PD 03-OCT-2002.
XX PF 27-MAR-2002; 2002WO-IB002163.
XX PR 27-MAR-2001; 2001GB-00007658.
XX PA (CHIR-) CHIRON SPA.
XX PA (GENO-) INST GENOMIC RES.
XX PI Masignani V, Tettelin H, Fraser C;
```

```
XX DR WPI; 2003-040579/03.
XX N-PSDB; ABX08011.
XX PT New proteins and nucleic acid molecules from Streptococcus pneumoniae,
XX PT useful as medicaments for treating or preventing a disease or infection
XX PT due to streptococcus bacteria, such as pneumonia, sepsis, otitis media or
XX PT ear infection.
XX PS Claim 1; SEQ ID NO 4598; 56pp; English.
XX
CC The invention relates to a protein comprising or having at least 50%
CC identity to any of the 2469 amino acid sequences, identified in the
CC specification (available on a computer readable format), or its fragment,
CC expressed from 2469 of 2489 identified DNA coding regions from the
CC Streptococcus pneumoniae type 4 strain genomic sequence appearing as
CC ABS56454. Also included are an antibody which binds one of the proteins,
CC treating a patient by administering the protein, DNA or antibody (in a
CC composition), a kit comprising first and second primers, which are the
CC nucleic acid cited above or fragments between nucleotides 8-100 of a
CC sequence not defined in the specification, for amplifying a target
CC sequence contained within a Streptococcus nucleic acid sequence, where
CC the first primer is substantially complementary to the target sequence
CC and the second primer is substantially complementary to the complement of
CC the target sequence, and where the parts of the primers having
CC substantial complementarity define the termini of the target sequence to
CC be amplified, assay comprising contacting a test compound with the
CC protein, and determining whether the test compound binds to the protein
CC and a Streptococcus pneumoniae bacterium, where one or more genes
CC encoding the proteins has been rendered inactive. The proteins, nucleic
CC acid molecules, antibody and compositions are useful as medicaments for
CC treating or preventing a disease or infection due to streptococcus
CC bacteria, particularly S. pneumoniae, such as pneumonia, sepsis, otitis
CC media or ear infection. They are also useful in developing vaccines,
CC diagnostics and antibiotics. The methods are useful for identifying
CC immunodominant proteins. The present sequence is one of the 2469 proteins
CC expressed by the identified coding regions from the genomic sequence.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences. (Updated on 23-OCT-2003 to
CC standardise OS field)
XX SQ Sequence 693 AA;
```

```
Query Match 100.0%; Score 2546; DB 6; Length 693;
Best Local Similarity 100.0%; Pred. No. 1.8e-152;
Matches 513; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MFASKSERKVHYSIRKFSVGASVVVASLVMSGVVHATENEGATQVPTSSNRANESQAEQ 60
Db 1 MFASKSERKVHYSIRKFSVGASVVVASLVMSGVVHATENEGATQVPTSSNRANESQAEQ 60

Qy 61 GEOPKLDSDRKARKEVEYVKIVGESYAKSTKKRHTITVALVNLNINKNEYLNKIV 120
Db 61 GEOPKLDSDRKARKEVEYVKIVGESYAKSTKKRHTITVALVNLNINKNEYLNKIV 120

Qy 121 ESTSESQQLILMMESRSKVDPAVSKFPEKDSOSSSSDSTKPEASDTAKPNKPTPEGEKV 180
Db 121 ESTSESQQLILMMESRSKVDPAVSKFPEKDSOSSSSDSTKPEASDTAKPNKPTPEGEKV 180

Qy 181 AEAKKVVEAEKKAQKQKEEDRRNYPITYTKTLEIAESDVEVKKAELVLVKVANEPR 240
Db 181 AEAKKVVEAEKKAQKQKEEDRRNYPITYTKTLEIAESDVEVKKAELVLVKVANEPR 240

Qy 241 DEQIKIQAEAVESKQAEATRLKKIKTDREAEAEAKRRADAKSQGPKGRKRGVPGEL 300
Db 241 DEQIKIQAEAVESKQAEATRLKKIKTDREAEAEAKRRADAKSQGPKGRKRGVPGEL 300

Qy 301 ATPDKKENDAKSSDSSVGEETLPSPLKPEKKVAEAEKKVVEAKKKAEDQKEEDRRNYP 360
Db 301 ATPDKKENDAKSSDSSVGEETLPSPLKPEKKVAEAEKKVVEAKKKAEDQKEEDRRNYP 360

Qy 361 NTYKTLLEIAESDVEVKKAELVLKVEAKPEPRNEEKVKQAKAVESKKAETRLKIKT 420
```



```

Db 361 NYKTLELEIAESDVEVKAELELVKEAEKPRNEEKVKQAEVSKAEATRLKIKT 420
Qy 421 DRKKAEEBAKKAABEDVKKEKPAQOPAPAPAKPAKPAKPENPAEQPKAEKPADQQ 480
Db 421 DRKKAEEBAKKAABEDVKKEKPAQOPAPAPAKPAKPAKPENPAEQPKAEKPADQQ 480
Qy 481 AEEDYARSEBEYNRLTQQPPKTEKPAQPSTP 513
Db 481 AEEDYARSEBEYNRLTQQPPKTEKPAQPSTP 513

RESULT 3
ADM92271
ID: ADM92271 standard; protein; 693 AA.
AC ADM92271;
XX
XX 03-JUN-2004 (first entry)
XX
XX S pneumoniae antigenic protein sequence SeqID468.
DE antibacterial; gene therapy; Streptococcus pneumoniae infection;
KW antigenic.
XX
XX Streptococcus pneumoniae.
OS
XX WO2004020609-A2.
XX
XX 11-MAR-2004.
XX
XX 02-SEP-2003; 2003WO-US027401.
XX
XX 30-AUG-2002; 2002US-0407082P.
XX
XX (TUFT ) UNIV TUFTS.
XX
XX Camilli A, Hava DL;
XX
XX N-PSDB; ADM92034.
XX
XX New Streptococcus pneumoniae nucleic acid molecules, useful for
XX diagnosing, treating and preventing active infections of Streptococcus
XX pneumoniae.
XX
XX Claim 27; SEQ ID NO 468; 123pp; English.
XX
XX This invention relates to novel isolated Streptococcus pneumoniae nucleic
XX acid molecules and the antigenic polypeptides encoded by them. The
XX invention may be useful for the production of compounds with an
XX antibacterial activity or for gene therapy. The nucleic acid molecules,
XX compositions and methods disclosed are useful for treating Streptococcus
XX pneumoniae infection. The present sequence is that of an S pneumoniae
XX protein of the invention.
XX
XX Sequence 693 AA;
XX
XX Query Match 100.0%; Score 2546; DB 8; Length 693;
XX Best Local Similarity 100.0%; Pred. No. 1.8e-152;
XX Matches 513; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 MFASKSERKVHYSIRKFSVGVSASVILVMSGVVHATENEGATQVPTSSNRANESQAEQ 60
XX
XX 1 MFASKSERKVHYSIRKFSVGVSASVILVMSGVVHATENEGATQVPTSSNRANESQAEQ 60
XX
XX 61 GEQPKLDSERDKARKEVEEYVKTVIGESYAKSTKKRHTITVALVNLNINKEVYNKIV 120
XX
XX 61 GEQPKLDSERDKARKEVEEYVKTVIGESYAKSTKKRHTITVALVNLNINKEVYNKIV 120
XX
XX 121 ESTSESQILMMESRSKVDEAVSKFEKDSSTSSSSSSSTKPEASDTAKPNKPTPEGKV 180
XX
XX 121 ESTSESQILMMESRSKVDEAVSKFEKDSSTSSSSSSSTKPEASDTAKPNKPTPEGKV 180

```

```

Qy 181 AEAKKKVEAEAKKADQKEEDRRNPTTITYKTLELEIAESDVEVKAELELVKVKANEPR 240
Db 181 AEAKKKVEAEAKKADQKEEDRRNPTTITYKTLELEIAESDVEVKAELELVKVKANEPR 240
Qy 241 DEOKTKQAEAEVSKQAEATRLKKIKTDREAEAEKERRADAKGKPKGRAKRGVPGEL 300
Db 241 DEOKTKQAEAEVSKQAEATRLKKIKTDREAEAEKERRADAKGKPKGRAKRGVPGEL 300
Qy 301 ATPDKKENDAKSSDSSVGEETLPSPLKPEKKVAAEAKKVEAEKKAEDQKEEDRRNPT 360
Db 301 ATPDKKENDAKSSDSSVGEETLPSPLKPEKKVAAEAKKVEAEKKAEDQKEEDRRNPT 360
Qy 361 NTYKTLELEIAESDVEVKAELELVKEAEKPRNEEKVKQAEVSKAEATRLKIKT 420
Db 361 NTYKTLELEIAESDVEVKAELELVKEAEKPRNEEKVKQAEVSKAEATRLKIKT 420
Qy 421 DRKKAEEBAKKAABEDVKKEKPAQOPAPAPAKPAKPAKPENPAEQPKAEKPADQQ 480
Db 421 DRKKAEEBAKKAABEDVKKEKPAQOPAPAPAKPAKPAKPENPAEQPKAEKPADQQ 480
Qy 481 AEEDYARSEBEYNRLTQQPPKTEKPAQPSTP 513
Db 481 AEEDYARSEBEYNRLTQQPPKTEKPAQPSTP 513

RESULT 4
AAY81653
ID AAY81653 standard; protein; 694 AA.
XX
XX AAY81653;
XX
XX 24-MAY-2000 (first entry)
XX
XX Streptococcus pneumoniae protein sequence ID302.
XX
XX Streptococcus pneumoniae; vaccine; screening; protein antigen;
XX antibacterial; antiinflammatory; meningitis; infection; diagnosis;
XX pneumococcal disease.
XX
XX Streptococcus pneumoniae.
XX
XX WO200006737-A2.
XX
XX 10-FEB-2000.
XX
XX 27-JUL-1999; 99WO-GB002451.
XX
XX 27-JUL-1998; 98GB-00016337.
XX
XX 19-MAR-1999; 99US-0125164P.
XX
XX (MICR-) MICROBIAL TECHNIQS LTD.
XX
XX Gilbert CFG, Hansbro PM;
XX
XX WPI; 2000-195300/17.
XX
XX New Streptococcal protein, useful as a vaccine, for diagnosis of
XX pneumococcal diseases and for screening agents capable of antagonizing or
XX inhibiting expression of the protein.
XX
XX Claim 2; Page 96; 108pp; English.
XX
XX AAY81501 to AAY81679 represent specifically claimed protein sequences
XX isolated from Streptococcus pneumoniae. AA05407 to AA05590 represent
XX specifically claimed nucleotide sequences isolated from S. pneumoniae.
XX The sequences have antibacterial and antiinflammatory properties. The
XX protein sequences, and fragments of them, are useful as immunogens and/or
XX antigens. The nucleotide sequences can be used in vaccines and in
XX diagnostic assays. The proteins and nucleotides can be useful for the
XX detection and diagnosis of S. pneumoniae. The protein sequences are also
XX useful for screening an agent capable of antagonising, inhibiting or
XX interfering with the function or expression of the proteins in which the

```

CC agent is useful for treatment or prophylaxis of *S. pneumoniae* infection
CC and meningitis. AAA05591 to AAA05614 represent primers used in the
CC exemplification of the present invention
XX
SQ Sequence 694 AA;

Query Match 100.0%; Score 2546; DB 3; Length 694;
Best Local Similarity 100.0%; Pred. No. 1.8e-152;
Matches 513; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MFASKSERKHYSTRKSGVSVASVWVSLVMSVGHATENEGATQVPTSSNRANESQAEQ 60
DB 1 MFASKSERKHYSTRKSGVSVASVWVSLVMSVGHATENEGATQVPTSSNRANESQAEQ 60
QY 61 GEQPKLDSERDKARKEVEEVKKIVGESYAKSTKKRHTITVALVNLNINKNYLNKIV 120
DB 61 GEQPKLDSERDKARKEVEEVKKIVGESYAKSTKKRHTITVALVNLNINKNYLNKIV 120
QY 121 ESTSESQILMMESRSKVDKAVSKFEKDSSSSSSSSTKPEASDTAKPNKPTPEGEKV 180
DB 121 ESTSESQILMMESRSKVDKAVSKFEKDSSSSSSSSTKPEASDTAKPNKPTPEGEKV 180
QY 181 AEAKKVEEAKKADQKEEDRRNYPTITYKTLEIAESDVKKAELELVKKANEPR 240
DB 181 AEAKKVEEAKKADQKEEDRRNYPTITYKTLEIAESDVKKAELELVKKANEPR 240
QY 241 DEQIKIQAEEVESKQAEATRLKKIKTDREAEAEAKRRADAKSQGKPKGRAKGVGPEL 300
DB 241 DEQIKIQAEEVESKQAEATRLKKIKTDREAEAEAKRRADAKSQGKPKGRAKGVGPEL 300
QY 301 ATPDKENDAKSSDSSVGEETLPSPLKPEKKVAAEKKVEAEKKAEADQKEEDRRNYPT 360
DB 301 ATPDKENDAKSSDSSVGEETLPSPLKPEKKVAAEKKVEAEKKAEADQKEEDRRNYPT 360
QY 361 NTYKTLEIAESDVKKAELELVKBEAKPRNEEKVKQAEEVESKQAEATRLKIKT 420
DB 361 NTYKTLEIAESDVKKAELELVKBEAKPRNEEKVKQAEEVESKQAEATRLKIKT 420
QY 421 DRKAAEAEAKKAAEEDKVEKPAEQOPAPAPKAEKPAPAPKAEKPAPAPKAEKPADQ 480
DB 421 DRKAAEAEAKKAAEEDKVEKPAEQOPAPAPKAEKPAPAPKAEKPAPAPKAEKPADQ 480
QY 481 AEEDYARRSEBEYNRLTQQQPPKTEKPAQSTP 513
DB 481 AEEDYARRSEBEYNRLTQQQPPKTEKPAQSTP 513

RESULT 5
AAAY32099
ID AAY32099 standard; protein; 655 AA.

AC AAY32099;

XX 01-FEB-2000 (first entry)

DE Choline binding protein A (CbpA).

KW Choline binding protein; CbpA; adhesin; immunogen; vaccine; diagnosis;
KW therapy; pneumococcus; sepsis; meningitis; otitis media; pneumonia.

OS Streptococcus pneumoniae.

FN WO9951188-A2.

PD 14-OCT-1999.

PF 07-APR-1999; 99WO-US007669.

PR 07-APR-1998; 98US-00056019.

PR 07-APR-1998; 98US-0080878P.

XX (SUUD-) ST JUDE CHILDREN'S RES HOSPITAL.

FA (MEDI-) MEDIMUNE INC.

XX
PI Tuomanen EI, Masure HR, Wizemann TM, Johnson LS, Koenig S;
XX WPI; 1999-633690/54.
DR N-PSDB; AAZ34452.
DR
XX New N-terminal choline binding protein A truncate polypeptides, used to
PT develop products for the diagnosis, prevention and treatment of
PT pneumococcal infections.
XX
PS Disclosure; Page 88-93; 160pp; English.

XX This sequence represents the choline binding protein A (CbpA) of
CC Streptococcus pneumoniae serotype type 4. The invention provides novel N-
CC terminal CbpA truncated polypeptides (see AAY32098-110 and AAY32179-90)
CC and polynucleotides encoding them, host-vector systems, and antibodies
CC that specifically bind to the polypeptides. The invention also relates to
CC vaccines including the polypeptides, which provide protection or elicit
CC protective antibodies to bacterial infection, specifically pneumococcus,
CC and to antibodies and antagonists against such polypeptides for use in
CC diagnosis and passive immunotherapy. The polypeptides and/or
CC polynucleotides are also useful as competitive inhibitors of bacterial
CC adhesion of pneumococcus
XX
SQ Sequence 655 AA;

Query Match 92.9%; Score 2366; DB 2; Length 655;
Best Local Similarity 100.0%; Pred. No. 4e-141;
Matches 475; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 39 ENEGATQVPTSSNRANESQAEQGEQPKLDSEDRDKARKEVEEVKKIVGESYAKSTKKRH 98
DB 1 ENEGATQVPTSSNRANESQAEQGEQPKLDSEDRDKARKEVEEVKKIVGESYAKSTKKRH 60
QY 99 TITVALVNLNINKNYLNKIVESTSESQILMMESRSKVDKAVSKFEKDSSSSSSDS 158
DB 61 TITVALVNLNINKNYLNKIVESTSESQILMMESRSKVDKAVSKFEKDSSSSSSDS 120
QY 159 STKPEASDTAKPNKPTPEGEKVAAEKKVBEAEKKADQKEEDRRNYPTITYKTLEIEIA 218
DB 121 STKPEASDTAKPNKPTPEGEKVAAEKKVBEAEKKADQKEEDRRNYPTITYKTLEIEIA 180
QY 219 ESDVEVKKAELELVKKANEPRDQIKIQAEEVESKQAEATRLKKIKTDREAEAEAKR 278
DB 181 ESDVEVKKAELELVKKANEPRDQIKIQAEEVESKQAEATRLKKIKTDREAEAEAKR 240
QY 279 RADAKEQKPKGRAKGVGPELATPDKKENDAKSSDSSVGEETLPSPLKPEKKVAAEAK 338
DB 241 RADAKEQKPKGRAKGVGPELATPDKKENDAKSSDSSVGEETLPSPLKPEKKVAAEAK 300
QY 339 KVEEAKKKAEDQKEEDRRNYPTNTYKTLEIEIAESDVKKAELELVKKEEAKPRNEEKV 398
DB 301 KVEEAKKKAEDQKEEDRRNYPTNTYKTLEIEIAESDVKKAELELVKKEEAKPRNEEKV 360
QY 399 KQAEAVESKQAEATRLKIKTDREAEAEAKRKAABEDKVEKPAEQOPAPAPKAEKP 458
DB 361 KQAEAVESKQAEATRLKIKTDREAEAEAKRKAABEDKVEKPAEQOPAPAPKAEKP 420
QY 459 APAPKPNPAEQKAEKPADQQAEDYARRSEBEYNRLTQQQPPKTEKPAQSTP 513
DB 421 APAPKPNPAEQKAEKPADQQAEDYARRSEBEYNRLTQQQPPKTEKPAQSTP 475

RESULT 6
AAAY49226
ID AAY49226 standard; protein; 655 AA.

XX AAAY49226;

XX 07-FEB-2000 (first entry)

DE CbpA of serotype 4 amino acid sequence.

XX

KW Choline binding protein A; CbpA; truncate; immune response; infection;
KW pneumococcal bacterium; vaccine.
OS Streptococcus pneumoniae.
XX WO9951187-A2.
XX PD 14-OCT-1999:
XX PF 07-APR-1999; 99WO-US007669.
XX PR 07-APR-1998; 98US-00056019.
XX PR 07-APR-1998; 98US-0080878P.
XX (SJUD-) ST JUDE CHILDREN'S RES HOSPITAL.
XX Tuomanen EI, Masure HR;
XX WPI; 1999-620161/53.
XX DR N-PSDB; AAZ31402.
XX Novel polypeptides, used to develop products for the diagnosis,
XX prevention and treatment of pneumococcal infections.
XX PS Disclosure; Page 9; 85pp; English.
XX The invention provides novel isolated polypeptides comprising the amino
XX acid sequence of an N-terminal choline binding protein A (CbpA) truncate
XX (CbpAT). The polypeptides can be selected from sequences shown in
XX CC AAY49225, AAY49227, AAY49230, AAY49231, AAY49233 and AAY49238. The
XX polypeptides can be used for inducing an immune response in a subject
XX which has been exposed to or infected with a pneumococcal bacterium. They
XX can also be used for preventing infection by a pneumococcal bacterium.
XX CC Vaccines comprising the polypeptides or encoding nucleic acids can be
XX used for treating a subject infected with or exposed to a pneumococcal
XX bacterium. Antibodies specifically binding the polypeptides can be used
XX for detection and diagnosis and for preventing pneumococcal attachment to
XX a mucosal surface. The products can be used in humans and other animals
XX such as domestic animals, such as feline or canine subjects, farm animals
XX such as bovine, equine, caprine, ovine, and porcine subjects, wild
XX animals (whether in the wild or in a zoological garden), research
XX animals, such as mice, rats, rabbits, goats, sheep, pigs, dogs, cats,
XX i.e. for veterinary medical use
SQ Sequence 655 AA;
Query Match 92.9%; Score 2366; DB 2; Length 655;
Best Local Similarity 100.0%; Pred. No. 4e-141;
Matches 475; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 39 ENEGATQVPTSSNRANESQAEQGEQPKLDSERDKARKEVEYVKKIIGESYAKSTKKRH 98
Db 1 ENEGATQVPTSSNRANESQAEQGEQPKLDSERDKARKEVEYVKKIIGESYAKSTKKRH 60
QY 99 TITVALNELNINIKNEYLNKIVESTSESQQLTLMMSRSKVDKAVKFEKSSSSSDS 158
Db 61 TITVALNELNINIKNEYLNKIVESTSESQQLTLMMSRSKVDKAVKFEKSSSSSDS 120
QY 159 STKPEASTAKPNKTEPEGEKVAEAKKVEEAKKQKQEDRRNYPTITTKTLELEIA 218
Db 121 STKPEASTAKPNKTEPEGEKVAEAKKVEEAKKQKQEDRRNYPTITTKTLELEIA 180
QY 219 ESDVEVKKAELELVKANEPRDEQIKQAEAEVESKQAEATRLKKIKTDREAEAEAKR 278
Db 181 ESDVEVKKAELELVKANEPRDEQIKQAEAEVESKQAEATRLKKIKTDREAEAEAKR 240
QY 279 RADAKEQKPKGRKRGVPGELATPDKKENDAKSSDSSVGEETLPSLPKPEKVAEAK 338
Db 241 RADAKEQKPKGRKRGVPGELATPDKKENDAKSSDSSVGEETLPSLPKPEKVAEAK 300
QY 339 KVEEAKKAEDQKEEDRRNYPTNTYKTLELEIAESDVEVKKAELELVKEEAKPEPNEEKV 398
Db 301 KVEEAKKAEDQKEEDRRNYPTNTYKTLELEIAESDVEVKKAELELVKEEAKPEPNEEKV 360

QY 399 KQAKAEVESKKAETRLLEKIKTDRKKAEEBAKRAABEDKVKKEPAEQQPAPAPKAEKP 458
Db 361 KQAKAEVESKKAETRLLEKIKTDRKKAEEBAKRAABEDKVKKEPAEQQPAPAPKAEKP 420
QY 459 APAPKPNPAEQPKAEKPADQQAEDYARRSEEEYNRLTQQQPPKTEKPAQPSSTP 513
Db 421 APAPKPNPAEQPKAEKPADQQAEDYARRSEEEYNRLTQQQPPKTEKPAQPSSTP 475
RESULT 7
ABW01575
ID ABW01575 standard; protein; 655 AA.
XX AC ABW01575;
XX 12-FEB-2004 (first entry)
XX Streptococcus pneumoniae cbpA serotype 4 protein.
XX Therapy; choline-binding protein A; cbpA; immune response; immunisation;
XX therapeutic; diagnosis; vaccine; antibacterial.
XX Streptococcus pneumoniae.
XX FH Key Location/Qualifiers
FT Misc-difference 344
FT /note= "Encoded by GCTT"
XX US2003096950-A1.
XX 22-MAY-2003.
XX 07-APR-1998; 98US-00056019.
XX 07-APR-1998; 98US-00056019.
XX (TUOM/) TUOMANEN E I.
XX (WIZE/) WIZEMANN T M.
XX (MASU/) MASURE H R.
XX (JOHN/) JOHNSON L S.
XX (KOEN/) KOENIG S.
XX Tuomanen EI, Wizemann TM, Masure HR, Johnson LS, Koenig S;
XX WPI; 2003-843081/78.
XX N-PSDB; AAD62881.
XX New N-terminal fragment of choline-binding protein A, useful for
XX treatment or prevention of infection by Streptococcus pneumoniae, also
XX new nucleic acid.
XX PS Disclosure; Page 3; Opp; English.
XX The present invention provides novel isolated polypeptide comprising an
XX amino acid sequence of an N-terminal choline-binding protein A (cbpA)
XX truncate. The invention is useful for induction of a protective or
XX therapeutic immune response against pneumococci (or other cross-reactive
XX bacteria). Antibodies raised against pneumococci are useful for passive
XX immunisation and diagnosis. The invention is also useful as vaccines. The
XX present sequence is Streptococcus pneumoniae choline-binding protein A
XX (cbpA) serotype 4 protein
SQ Sequence 655 AA;
Query Match 92.9%; Score 2366; DB 7; Length 655;
Best Local Similarity 100.0%; Pred. No. 4e-141;
Matches 475; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 39 ENEGATQVPTSSNRANESQAEQGEQPKLDSERDKARKEVEYVKKIIGESYAKSTKKRH 98
Db 1 ENEGATQVPTSSNRANESQAEQGEQPKLDSERDKARKEVEYVKKIIGESYAKSTKKRH 60

Qy	99	TIITVALVNELNNIKNEYLINKIVESTSESQOLILMMESRSKVDEAVSKPEKSDSSSSSSSDS	158
Db	61	TIITVALVNELNNIKNEYLINKIVESTSESQOLILMMESRSKVDEAVSKPEKSDSSSSSSSDS	120
Qy	159	STKPEASDTAKPNKPTEPGKEKVABAAKKVBEAAKKAQKQEBDRNYPITIYKTLLELETA	218
Db	121	STKPEASDTAKPNKPTEPGKEKVABAAKKVBEAAKKAQKQEBDRNYPITIYKTLLELETA	180
Qy	219	ESDVEVVKAAELVLVKVKANEPREDEQIKQAAEAESVKQAAEATRLKKIKTDRBEAEAEAKR	278
Db	181	ESDVEVVKAAELVLVKVKANEPREDEQIKQAAEAESVKQAAEATRLKKIKTDRBEAEAEAKR	240
Qy	279	RADAKEQKPKGRAGRVGPGELATPDKKENDAKSSDSSVGBETLPSPLKPKKVABAEAK	338
Db	241	RADAKEQKPKGRAGRVGPGELATPDKKENDAKSSDSSVGBETLPSPLKPKKVABAEAK	300
Qy	339	KVEBAKKAQEQKEEDRRNYPNTIYKTLLELETAESDVEVKAAELVLKVEAKEPRNEKV	398
Db	301	KVEBAKKAQEQKEEDRRNYPNTIYKTLLELETAESDVEVKAAELVLKVEAKEPRNEKV	360
Qy	399	KQAAEAESKKAAEATRLLEKIKTDRKAAEEAAKRAAEEDVKKEPAEQPOPAPAPKAEKP	458
Db	361	KQAAEAESKKAAEATRLLEKIKTDRKAAEEAAKRAAEEDVKKEPAEQPOPAPAPKAEKP	420
Qy	459	APAKPENPAPQPKAEKPADQQAABEDYARRSEEEYNRLTQQOOPKTEKPAQBPSTP	513
Db	421	APAKPENPAPQPKAEKPADQQAABEDYARRSEEEYNRLTQQOOPKTEKPAQBPSTP	475

RESULT 8	
ADRI0795	
ID	ADRI0795 standard; protein; 655 AA.
XX	
XX	ADRI0795;
XX	
XX	23-SEP-2004 (first entry)
XX	
XX	Choline binding protein A (Cbpa) serotype type 4.
XX	
XX	vaccine; pneumococcal infection; Streptococcus pneumoniae;
KW	choline binding protein A; Cbpa.
KW	
XX	
OS	Streptococcus pneumoniae.
XX	
XX	
Key	Location/Qualifiers
PH	
FT	Misc-difference 344
FT	/note= "Encoded by GCTT"

	Query Match	Score	DB	Length
CC	pharmaceutical carrier where the polypeptide does not bind to choline, CC exhibits a tertiary structure as found in a native, full-length CBPA polypeptide, and the polypeptide content of the vaccine is for treating CC protecting against pneumococcal infection. The vaccine is useful for CC treating or protecting against pneumococcal infection, by inhibiting the CC adhesion of <i>Pneumococcus</i> , particularly <i>Streptococcus pneumoniae</i> . The CC present sequence represents the amino acid sequence of choline binding CC protein A (Cbpa) serotype type 4.	92.9%;	2366;	DB 8; Length 655;
XX	Best Local Similarity	100.0%;	Pred. No. 4e-141;	
SQ	Matches	475;	Conservative	0; Mismatches
		0;	Indels	0; Gaps
QY	39	ENEGATQVPTSSNNRANESQAEQGGQPKLDSERDKARKEVEEYVKKI	VGESYAKSTKKRH	98
DB	1	ENEGATQVPTSSNNRANESQAEQGGQPKLDSERDKARKEVEEYVKKI	VGESYAKSTKKRH	60
QY	99	TITVALVNELNNIKNEYLNKIVSTESBSQQLMMESRSKVDEAVSKPEKSSSSSSSDS	158	
DB	61	TITVALVNELNNIKNEYLNKIVSTESBSQQLMMESRSKVDEAVSKPEKSSSSSSSDS	120	
QY	159	STKPEASDTAKPNKPTPEGKVAEAKKKVBEAEKKAQDKQEDRRNYPTIITKTLLELEIA	218	
DB	121	STKPEASDTAKPNKPTPEGKVAEAKKKVBEAEKKAQDKQEDRRNYPTIITKTLLELEIA	180	
QY	219	ESDVEVKKAELVVKVXANEPDRDQKIKQAEAEVSKQAEATRLKKIKTDREAEAEAAKR	278	
DB	181	ESDVEVKKAELVVKVXANEPDRDQKIKQAEAEVSKQAEATRLKKIKTDREAEAEAAKR	240	
QY	279	RADAEQGGPKGRAKRGVGPGLATPDKKENDAKSSDSVSGBETI	PSPSLKPEKKVAAEAK	338
DB	241	RADAEQGGPKGRAKRGVGPGLATPDKKENDAKSSDSVSGBETI	PSPSLKPEKKVAAEAK	300
QY	339	KVEEAKKKAEDQKEEDRRNYPTNTYKTLELEIAESDVEVKKAELVKAEAKEPNEEKV	398	
DB	301	KVEEAKKKAEDQKEEDRRNYPTNTYKTLELEIAESDVEVKKAELVKAEAKEPNEEKV	360	
QY	399	KQAKAEVSKKAEATRLLEKIKTDKKAEEAAKRAAEEDKVKEPAEQPQAPAPKAEKP	458	
DB	361	KQAKAEVSKKAEATRLLEKIKTDKKAEEAAKRAAEEDKVKEPAEQPQAPAPKAEKP	420	
QY	459	APAPKPNPAEQPKAEKPADQQAEDYARSEEBEYNRLTQQOPKTEKPAQPSPT	513	
DB	421	APAPKPNPAEQPKAEKPADQQAEDYARSEEBEYNRLTQQOPKTEKPAQPSPT	475	

RESULT 9
 AAY49250
 ID AAY49250 standard; protein; 460 AA.
 XX
 AC AAY49250;
 XX
 XX
 DT 07-FEB-2000 (first entry)
 XX
 DE N-terminal region of CbpA polypeptide Ntype4.
 XX
 KW Choline binding protein A; CbpA; truncate; immune response; infection;
 KW pneumococcal bacterium; vaccine.
 KW

XX
PI Tuomanen EI, Masure HR;
XX WPI; 1999-620161/53.
XX Novel polypeptides, used to develop products for the diagnosis,
PT prevention and treatment of pneumococcal infections.
XX
XX Disclosure; Fig 2A-B; 85pp; English.
XX
XX The invention provides novel isolated polypeptides comprising the amino
CC acid sequence of an N-terminal choline binding protein A (Cbpa) truncate
CC (CbpaT). The polypeptides can be selected from sequences shown in
CC AAY49225, AAY49227, AAY49230, AAY49231, AAY49233 and AAY49238. The
CC polypeptides can be used for inducing an immune response in a subject
CC which has been exposed to or infected with a pneumococcal bacterium. They
CC can also be used for preventing infection by a pneumococcal bacterium.
CC Vaccines comprising the polypeptides or encoding nucleic acids can be
CC used for treating a subject infected with or exposed to a pneumococcal
CC bacterium. Antibodies specifically binding the polypeptides can be used
CC for detection and diagnosis and for preventing pneumococcal attachment to
CC a mucosal surface. The products can be used in humans and other animals
CC such as domestic animals, such as feline or canine subjects, farm animals
CC such as bovine, equine, caprine, ovine, and porcine subjects, wild
CC animals (whether in the wild or in a zoological garden), research
CC animals, such as mice, rats, rabbits, goats, sheep, pigs, dogs, cats,
CC i.e. for veterinary medical use
XX
XX Sequence 460 AA;
Query Match 89.1%; Score 2268; DB 2; Length 460;
Best Local Similarity 99.8%; Pred. No. 4.1e-135;
Matches 458; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 25 VVASLWVGSVVHATENEGATOVPTSSNRANESQAEGQPKLDSERDKARKEVEYVKK 84
DB 2 IVASLWVGSVVHATENEGATOVPTSSNRANESQAEGQPKLDSERDKARKEVEYVKK 61
QY 85 IVGESYAKSTKKRHTITVALVNLNINIKVEYLNKIVESTSESQILMWESRSKYDEAVS 144
DB 62 IVGESYAKSTKKRHTITVALVNLNINIKVEYLNKIVESTSESQILMWESRSKYDEAVS 121
QY 145 KFEKDSRSSSSSSDSTKPEASDTAKPNKPTPEGKVAEAKKKVBEAEKAKDQKEEDRN 204
DB 122 KFEKDSRSSSSSSDSTKPEASDTAKPNKPTPEGKVAEAKKKVBEAEKAKDQKEEDRN 181
QY 205 YPTITYKTLELEIAESDVEVKKAELELVKVKANEPDEQIKQAEAEVSKQAEATRLKK 264
DB 182 YPTITYKTLELEIAESDVEVKKAELELVKVKANEPDEQIKQAEAEVSKQAEATRLKK 241
QY 265 IKTDREAEAEAKRADAKAQKPKGRKRGVPGELATPDKKENDAKSSDSSVSGEETLPS 324
DB 242 IKTDREAEAEAKRADAKAQKPKGRKRGVPGELATPDKKENDAKSSDSSVSGEETLPS 301
QY 325 PSLKPEKKVAEAEKKEVBEAKKAEADQKEEDRRNPTNTYKTLELEIAESDVEVKKAELEL 384
DB 302 PSLKPEKKVAEAEKKEVBEAKKAEADQKEEDRRNPTNTYKTLELEIAESDVEVKKAELEL 361
QY 385 VKEAKEPRNEEKVKQAEAEVSKQAEATRLKIKTDREAEAEAKRADAKAQKPKGRKRGVPGELATPDKKENDAKSSDSSVSGEETLPS 444
DB 362 VKEAKEPRNEEKVKQAEAEVSKQAEATRLKIKTDREAEAEAKRADAKAQKPKGRKRGVPGELATPDKKENDAKSSDSSVSGEETLPS 421
QY 445 EPOQAPAPKAEKAPAPKPPENPAPQKAEKPADQQAEE 483
DB 422 EPOQAPAPKAEKAPAPKPPENPAPQKAEKPADQQAEE 460
RESULT 10
ADRI0831
ID ADRI0831 standard; protein; 460 AA.
XX
AC ADRI0831;
XX

DT 23-SEP-2004 (first entry)
XX Choline binding protein A (Cbpa) N-terminal region Ntype4.
DE
XX vaccine; pneumococcal infection; Streptococcus pneumoniae;
KW choline binding protein A; Cbpa.
KW Streptococcus pneumoniae.
OS
XX US2004120966-A1.
PN
XX 24-JUN-2004.
PD
XX 05-JAN-2004; 2004US-00751702.
PF
XX 07-APR-1998; 98US-00056019.
PR
XX (SUJD-) ST JUDE CHILDREN'S RES HOSPITAL.
PA (MEDI-) MEDIMUNE INC.
XX
XX Tuomanen EI, Witzemann TM, Masure HR, Johnson LS, Koenig S;
PI WPI; 2004-479616/45.
XX
XX New vaccine comprises a polypeptide, which does not bind to choline and
PT exhibits a tertiary structure as found in a native, full-length Cbpa
PT polypeptide, useful for treating or protecting against pneumococcal
PT infection.
XX
XX Disclosure; SEQ ID NO 38; 70pp; English.
PS
XX The invention relates to a vaccine for treating or protecting against
XX pneumococcal infection. The vaccine comprises a polypeptide in a
CC pharmaceutical carrier where the polypeptide does not bind to choline,
CC exhibits a tertiary structure as found in a native, full-length Cbpa
CC polypeptide, and the polypeptide content of the vaccine is for treating
CC or protecting against pneumococcal infection. The vaccine is useful for
CC treating or protecting against pneumococcal infection, by inhibiting the
CC adhesion of Pneumococcus, particularly Streptococcus pneumoniae. The
CC present sequence represents the amino acid sequence of a choline binding
XX protein A (Cbpa) N-terminal region.
XX Sequence 460 AA;
Query Match 89.1%; Score 2268; DB 8; Length 460;
Best Local Similarity 99.8%; Pred. No. 4.1e-135;
Matches 458; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 25 VVASLWVGSVVHATENEGATOVPTSSNRANESQAEGQPKLDSERDKARKEVEYVKK 84
DB 2 IVASLWVGSVVHATENEGATOVPTSSNRANESQAEGQPKLDSERDKARKEVEYVKK 61
QY 85 IVGESYAKSTKKRHTITVALVNLNINIKVEYLNKIVESTSESQILMWESRSKYDEAVS 144
DB 62 IVGESYAKSTKKRHTITVALVNLNINIKVEYLNKIVESTSESQILMWESRSKYDEAVS 121
QY 145 KFEKDSRSSSSSSDSTKPEASDTAKPNKPTPEGKVAEAKKKVBEAEKAKDQKEEDRN 204
DB 122 KFEKDSRSSSSSSDSTKPEASDTAKPNKPTPEGKVAEAKKKVBEAEKAKDQKEEDRN 181
QY 205 YPTITYKTLELEIAESDVEVKKAELELVKVKANEPDEQIKQAEAEVSKQAEATRLKK 264
DB 182 YPTITYKTLELEIAESDVEVKKAELELVKVKANEPDEQIKQAEAEVSKQAEATRLKK 241
QY 265 IKTDREAEAEAKRADAKAQKPKGRKRGVPGELATPDKKENDAKSSDSSVSGEETLPS 324
DB 242 IKTDREAEAEAKRADAKAQKPKGRKRGVPGELATPDKKENDAKSSDSSVSGEETLPS 301
QY 325 PSLKPEKKVAEAEKKEVBEAKKAEADQKEEDRRNPTNTYKTLELEIAESDVEVKKAELEL 384
DB 302 PSLKPEKKVAEAEKKEVBEAKKAEADQKEEDRRNPTNTYKTLELEIAESDVEVKKAELEL 361
QY 385 VKEAKEPRNEEKVKQAEAEVSKQAEATRLKIKTDREAEAEAKRADAKAQKPKGRKRGVPGELATPDKKENDAKSSDSSVSGEETLPS 444

```

Db      362 VKBEAKPRNEEKVKQAKAEVSKAEATRLLEKI KTDKKAEEBAKRAAEEDKVKSKPA 421
Qy      445 EQQPAPAPKAEPAPAPKPPENPAEQPKAEKPADQQAEE 483
Db      422 EQQPAPAPKAEPAPAPKPPENPAEQPKAEKPADQQAEE 460

RESULT 11
AAY32189
ID AAY32189 standard; protein; 460 AA.
AC AAY32189;
XX
XX
XX 01-FEB-2000 (first entry)
XX
XX N-terminal choline binding protein A (Cbpa) truncate.
XX
XX Choline binding protein; Cbpa; truncate; adhesin; immunogen; vaccine;
XX diagnosis; therapy; pneumococcus; sepsis; meningitis; otitis media;
XX pneumonia.
XX
XX Streptococcus pneumoniae.
XX
XX Synthetic.
XX
XX WO9951188-A2.
XX
XX 14-OCT-1999.
XX
XX 07-APR-1999; 99WO-US007669.
XX
XX 07-APR-1998; 98US-00056019.
XX 07-APR-1998; 98US-0080878P.
XX
XX (SUJUD-) ST JUDE CHILDREN'S RES HOSPITAL.
XX (MEDI-) MEDIMUNE INC.
XX
XX Tuomanen EI, Masure HR, Wizemann TM, Johnson LS, Koenig S;
XX WPI; 1999-633690/54.
XX
XX New N-terminal choline binding protein A truncate polypeptides, used to
XX develop products for the diagnosis, prevention and treatment of
XX pneumococcal infections.
XX
XX Claim 47; Fig 2A-B; 160pp; English.
XX
XX The present sequence represents an N-terminal choline binding protein A
XX (Cbpa) truncate, denoted Ntype 4 Cbpa trun, derived from Streptococcus
XX pneumoniae Cbpa. Claimed vaccines contain and N-terminal Cbpa truncate
XX such as the present sequence, or a polypeptide comprising a conserved
XX region of the Cbpa truncate. The vaccines provide protection or elicit
XX protective antibodies to bacterial infection, specifically pneumococcus.
XX Antibodies and antagonists against the N-terminal Cbpa truncates are
XX used in diagnosis and passive immunotherapy
XX
XX Sequence 460 AA;

Query Match      89.0%; Score 2265; DB 2; Length 460;
Best Local Similarity 99.6%; Pred. No. 6.4e-135;
Matches 457; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy      25 VVASLWVGSVVHATENEGATQVPTSSNRANESQAEQEQPKLDSERDKARKEVEYVKK 84
Db      2 IVASLWVGSVVHATENEGATQVPTSSNRANESQAEQEQPKLDSERDKARKEVEYVKK 61
Qy      85 IVGESYAKSTKKRHTITVALVNLNINIKNEYNLKNIVESTSESQILMMSRSKVDKAVS 144
Db      62 IVGESYAKSTKKRHTITVALVNLNINIKNEYNLKNIVESTSESQILMMSRSKVDKAVS 121
Qy      145 KFEKDSSSSSSDSTKPEASDTAKPNKPTPEGKVAEAKKKAEAKKADQKEEDRN 204
Db      122 KFEKDSSSSSSDSTKPEASDTAKPNKPTPEGKVAEAKKKAEAKKADQKEEDRN 181

```

```

Qy      205 YPTITYKTLELEIAESDVEVKAELELVKVKANEPDRDEQKIKQAEAEVESKQAEATRLKK 264
Db      182 YPTITYKTLELEIAESDVEVKAELELVKVKANEPDRDEQKIKQAEAEVESKQAEATRLKK 241
Qy      265 IKTDREBAEEAKRRADAKQOGKPKGRAKGVPGELATPDKKENDAKSSSSVGEETLPS 324
Db      242 IKTDREBAEEAKRRADAKQOGKPKGRAKGVPGELATPDKKENDAKSSSSVGEETLPS 301
Qy      325 PSLKPEKKVAEAKKVEEAKKAEADQKEEDRRNYPNTYKTLELEIAESDVEVKAELEL 384
Db      302 PSLKPEKKVAEAKKVEEAKKAEADQKEEDRRNYPNTYKTLELEIAESDVEVKAELEL 361
Qy      385 VKBEAKPRNEEKVKQAKAEVSKAEATRLLEKI KTDKKAEEBAKRAAEEDKVKSKPA 444
Db      362 VKBEAKPRNEEKVKQAKAEVSKAEATRLLEKI KTDKKAEEBAKRAAEEDKVKSKPA 421
Qy      445 EQQPAPAPKAEPAPAPKPPENPAEQPKAEKPADQQAEE 483
Db      422 EQQPAPAPKAEPAPAPKPPENPAEQPKAEKPADQQAEE 460

RESULT 12
AAY32190
ID AAY32190 standard; protein; 459 AA.
XX
XX AAY32190;
XX
XX 01-FEB-2000 (first entry)
XX
XX N-terminal choline binding protein A (Cbpa) truncate.
XX
XX Choline binding protein; Cbpa; truncate; adhesin; immunogen; vaccine;
XX diagnosis; therapy; pneumococcus; sepsis; meningitis; otitis media;
XX pneumonia.
XX
XX Streptococcus pneumoniae.
XX
XX Synthetic.
XX
XX WO9951188-A2.
XX
XX 14-OCT-1999.
XX
XX 07-APR-1999; 99WO-US007669.
XX
XX 07-APR-1998; 98US-00056019.
XX 07-APR-1998; 98US-0080878P.
XX
XX (SUJUD-) ST JUDE CHILDREN'S RES HOSPITAL.
XX (MEDI-) MEDIMUNE INC.
XX
XX Tuomanen EI, Masure HR, Wizemann TM, Johnson LS, Koenig S;
XX WPI; 1999-633690/54.
XX
XX New N-terminal choline binding protein A truncate polypeptides, used to
XX develop products for the diagnosis, prevention and treatment of
XX pneumococcal infections.
XX
XX Claim 47; Fig 2A-B; 160pp; English.
XX
XX The present sequence represents an N-terminal choline binding protein A
XX (Cbpa) truncate, denoted ATCC4 Cbpa trun, derived from Streptococcus
XX pneumoniae Cbpa. Claimed vaccines contain and N-terminal Cbpa truncate
XX such as the present sequence, or a polypeptide comprising a conserved
XX region of the Cbpa truncate. The vaccines provide protection or elicit
XX protective antibodies to bacterial infection, specifically pneumococcus.
XX Antibodies and antagonists against the N-terminal Cbpa truncates are
XX used in diagnosis and passive immunotherapy
XX
XX Sequence 459 AA;

Query Match      88.4%; Score 2251; DB 2; Length 459;

```

Best Local Similarity 98.9%; Pred. No. 4.9e-134; Matches 454; Conservative 2; Mismatches 3; Indels 0; Gaps 0;	
QY 25	VVASLVMSGVVHATNEGATQVPTSSNRANESQAEGQPKKLDSEDKARKEVEEYVKK 84
Db 1	IVASLVMSGVVHATNEGATQVPTSSNRANESQAEGQPKKLDSEDKARKEVEEYVKK 60
QY 85	IVGESYAKSTKKRHTITVALVNLNIIKNEYLKIVESTSESQLOILMMESRSKYDEAVS 144
Db 61	IVGESYAKSTKKRHTITVALVNLNIIKNEYLKIVESTSESQLOILMMESRSKYDEAVS 120
QY 145	KFEKDSSSSSSDSTKPEASDTAKPNKPTPEGKVAEAKKVEAEKKAQKQEDRRN 204
Db 121	KFEKDSSSSSSDSTKPEASDTAKPNKPTPEGKVAEAKKVEAEKKAQKQEDRRN 180
QY 205	YPTITYKTLELEIAESDVEVKKAELELVKVKANEPDRDQKIQAQAEVSKQAEATRLKK 264
Db 181	YPTITYKTLELEIAESDVEVKKAELELVKVKANEPDRDQKIQAQAEVSKQAEATRLKK 240
QY 265	IKTDREAEAEAKRADAKEQKPKGRKRGVPGELATPDKKENDAKSSDSSVGETLPS 324
Db 241	IKTDREAEAEAKRADAKEQKPKGRKRGVPGELATPDKKENDAKSSDSSVGETLPS 300
QY 325	PSLKPEKKVAEAEKVEAEKKAQKQEDRRNYPNTYKTLELEIAESDVEVKKAELEL 384
Db 301	PSLKPEKKVAEAEKVEAEKKAQKQEDRRNYPNTYKTLELEIAESDVEVKKAELEL 360
QY 385	VKEAKEPRNEEKVKQAKAEVSKKAEATRLKIKTDRKKAEEAEAKRAAEEDKVKEKPA 444
Db 361	VKEAKEPRNEEKVKQAKAEVSKKAEATRLKIKTDRKKAEEAEAKRAAEEDKVKEKPA 420
QY 445	EQQPAPAPKAEKPAPAPKPNPAPQKAEKPADQQAEE 483
Db 421	EQQPAPAPKTEKPAPAPKPNPAPQKAEKPADQQAEE 459
RESULT 14	
ADRI0832	
ID ADRI0832	standard; protein; 459 AA.
XX AC	ADRI0832;
XX DT	23-SEP-2004 (first entry)
XX DE	Choline binding protein A (Cbpa) N-terminal region ATCC4.
XX KW	vaccine; pneumococcal infection; Streptococcus pneumoniae;
XX KW	choline binding protein A; Cbpa.
XX OS	Streptococcus pneumoniae.
XX PN	US2004120966-A1.
XX PD	24-JUN-2004.
RESULT 13	
AA49251	
ID AA49251	standard; protein; 459 AA.
XX AC	AA49251;
XX DT	07-FEB-2000 (first entry)
XX DE	N-terminal region of CbpA polypeptide ATCC4.
XX KW	Choline binding protein A; CbpA; truncate; immune response; infection;
XX KW	pneumococcal bacterium; vaccine.
XX OS	Streptococcus pneumoniae.
XX PN	WO9951187-A2.
XX PD	14-OCT-1999.
XX PF	07-APR-1999; 99WO-US007668.
XX PR	07-APR-1998; 98US-00056019.
XX PR	07-APR-1998; 98US-0080878P.
XX PA	(SJD-) ST JUDE CHILDREN'S RES HOSPITAL.
XX PI	Tuomanen EI, Masure HR;
XX DR	WPI; 1999-620161/53.
XX PT	Novel polypeptides, used to develop products for the diagnosis,
XX PT	prevention and treatment of pneumococcal infections.
XX PS	Disclosure; Fig 2A-B; 85pp; English.
XX CC	The invention provides novel isolated polypeptides comprising the amino
XX CC	acid sequence of an N-terminal choline binding protein A (Cbpa) truncate

(CbpaT). The polypeptides can be selected from sequences shown in CC AA49225, AA49227, AA49230, AA49231, AA49233 and AA49238. The CC polypeptides can be used for inducing an immune response in a subject CC which has been exposed to or infected with a pneumococcal bacterium. CC can also be used for preventing infection by a pneumococcal bacterium. CC Vaccines comprising the polypeptides or encoding nucleic acids can be CC used for treating a subject infected with or exposed to a pneumococcal CC bacterium. Antibodies specifically binding the polypeptides can be used CC for detection and diagnosis and for preventing pneumococcal attachment to CC a mucosal surface. The products can be used in humans and other animals CC such as domestic animals, such as feline or canine subjects, farm animals CC such as bovine, equine, caprine, ovine, and porcine subjects, wild CC animals (whether in the wild or in a zoological garden), research CC animals, such as mice, rats, rabbits, goats, sheep, pigs, dogs, cats, CC i.e. for veterinary medical use

XX Sequence 459 AA;

Query Match	88.4%;	Score 2251;	DB 2;	Length 459;
Best Local Similarity	98.9%;	Pred. No. 4.9e-134;		
Matches 454;	Conservative 2;	Mismatches 3;	Indels 0;	Gaps 0;

QY 25 VVASLVMSGVVHATNEGATQVPTSSNRANESQAEGQPKKLDSEDKARKEVEEYVKK 84
Db 1 IVASLVMSGVVHATNEGATQVPTSSNRANESQAEGQPKKLDSEDKARKEVEEYVKK 60
QY 85 IVGESYAKSTKKRHTITVALVNLNIIKNEYLKIVESTSESQLOILMMESRSKYDEAVS 144
Db 61 IVGESYAKSTKKRHTITVALVNLNIIKNEYLKIVESTSESQLOILMMESRSKYDEAVS 120
QY 145 KFEKDSSSSSSDSTKPEASDTAKPNKPTPEGKVAEAKKVEAEKKAQKQEDRRN 204
Db 121 KFEKDSSSSSSDSTKPEASDTAKPNKPTPEGKVAEAKKVEAEKKAQKQEDRRN 180
QY 205 YPTITYKTLELEIAESDVEVKKAELELVKVKANEPDRDQKIQAQAEVSKQAEATRLKK 264
Db 181 YPTITYKTLELEIAESDVEVKKAELELVKVKANEPDRDQKIQAQAEVSKQAEATRLKK 240
QY 265 IKTDREAEAEAKRADAKEQKPKGRKRGVPGELATPDKKENDAKSSDSSVGETLPS 324
Db 241 IKTDREAEAEAKRADAKEQKPKGRKRGVPGELATPDKKENDAKSSDSSVGETLPS 300
QY 325 PSLKPEKKVAEAEKVEAEKKAQKQEDRRNYPNTYKTLELEIAESDVEVKKAELEL 384
Db 301 PSLKPEKKVAEAEKVEAEKKAQKQEDRRNYPNTYKTLELEIAESDVEVKKAELEL 360
QY 385 VKEAKEPRNEEKVKQAKAEVSKKAEATRLKIKTDRKKAEEAEAKRAAEEDKVKEKPA 444
Db 361 VKEAKEPRNEEKVKQAKAEVSKKAEATRLKIKTDRKKAEEAEAKRAAEEDKVKEKPA 420
QY 445 EQQPAPAPKAEKPAPAPKPNPAPQKAEKPADQQAEE 483
Db 421 EQQPAPAPKTEKPAPAPKPNPAPQKAEKPADQQAEE 459

RESULT 14
ADRI0832
ID ADRI0832 standard; protein; 459 AA.
XX AC ADRI0832;
XX DT 23-SEP-2004 (first entry)
XX DE Choline binding protein A (Cbpa) N-terminal region ATCC4.
XX KW vaccine; pneumococcal infection; Streptococcus pneumoniae;
XX KW choline binding protein A; Cbpa.
XX OS Streptococcus pneumoniae.
XX PN US2004120966-A1.
XX PD 24-JUN-2004.

XX 05-JAN-2004; 2004US-00751702.
 XX 07-APR-1998; 98US-00056019.
 XX (SJUD-) ST JUDE CHILDREN'S RES HOSPITAL.
 PA (MEDI-) MEDIMUNE INC.
 XX Tuomanen EI, Wizemann TM, Masure HR, Johnson LS, Koenig S;
 PI WPI; 2004-479616/45.
 XX
 XX New vaccine comprises a polypeptide, which does not bind to choline and
 PT exhibits a tertiary structure as found in a native, full-length CbpA
 PT polypeptide, useful for treating or protecting against pneumococcal
 PT infection.
 XX
 XX Disclosure; SEQ ID NO 39; 70pp; English.
 XX
 XX The invention relates to a vaccine for treating or protecting against
 CC pneumococcal infection. The vaccine comprises a polypeptide in a
 CC pharmaceutical carrier where the polypeptide does not bind to choline,
 CC exhibits a tertiary structure as found in a native, full-length CbpA
 CC polypeptide, and the polypeptide content of the vaccine is for treating
 CC or protecting against pneumococcal infection. The vaccine is useful for
 CC treating or protecting against pneumococcal infection, by inhibiting the
 CC adhesion of Pneumococcus, particularly Streptococcus pneumoniae. The
 CC present sequence represents the amino acid sequence of a choline binding
 CC protein A (CbpA) N-terminal region.
 XX
 SQ Sequence 459 AA;

Query Match 88.4%; Score 2251; DB 8; Length 459;
 Best Local Similarity 98.9%; Pred. No. 4.9e-134;
 Matches 454; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
 QY 25 VVASLVGMSVVVHATENEGATQVPTSSNRANESQAEQEQPKLDSERDKARKEVEEVYVK 84
 :|||||
 DB 1 IVASLVGMSVVVHATENEGATQVPTSSNRANESQAEQEQPKLDSERDKARKEVEEVYVK 60
 QY 85 IVGESYAKSTKKRHTITVALVNLNINKNEYLKIVESTSESQILMMESRSKVDKAVS 144
 DB 61 IVGESYAKSTKKRHTITVALVNLNINKNEYLKIVESTSESQILMMESRSKVDKAVS 120
 QY 145 KFEKDSSSSSSDSTKPEASDTAKPNKPTPEGKVAEAKKKVEAEKAKDOKEEDRN 204
 DB 121 KFEKDSSSSSSDSTKPEASDTAKPNKPTPEGKVAEAKKKVEAEKAKDOKEEDRN 180
 QY 205 YPTITYKTLELEIAESDVEVKAELELVKVKANPRDEQKIQAEEVESKQAEATRLKK 264
 DB 181 YPTITYKTLELEIAESDVEVKAELELVKVKANPRDKQKIQAEEVESKQAEATRLKK 240
 QY 265 IKTDREAEBAKRADAKGQKPKGRAGVPGELATPDKKENDAKSSDSSVGEETLPS 324
 DB 241 IKTDREAEBAEKRAADAKGQKPKGRVPGELATPDKKENDAKSSDSSVGEETLPS 300
 QY 325 PSLKPEKKVAEAKKVEBAKKAKDOKEEDRRNPTNTYKTLELEIAESDVEVKAELEL 384
 DB 301 PSLKPEKKVAEAKKVEBAKKAKDOKEEDRRNPTNTYKTLELEIAESDVEVKAELEL 360
 QY 385 VKEAKEPRNEKVKQAKAEVESKKAETRLLEKIKTDKKAEEBAKKAFAEEDVKKEKPA 444
 DB 361 VKEAKEPRNEKVKQAKAEVESKKAETRLLEKIKTDKKAEEBAKKAFAEEDVKKEKPA 420
 QY 445 EQQPAPAPKAEKAPAPKPNPAEQKAEKPADQQAEE 483
 DB 421 EQQPAPAPKTEKAPAPKPNPAEQKAEKPADQQAEE 459

RESULT 15
 AAY49143
 ID AAY49143 standard; protein; 446 AA.
 XX

AC AAY49143;
 XX 17-JAN-2000 (first entry)
 XX Amino acid sequence of choline-binding protein fragment #7.
 DE Truncated surface binding protein; alpha helix; choline binding protein;
 XX vaccine; invasive bacterial infection; otitis media; sepsis; meningitis;
 KW lobar pneumonia infection; antibody; immature immune system;
 KW immunocompromised.
 XX Streptococcus pneumoniae.
 OS
 XX WO9951266-A2.
 FN 14-OCT-1999.
 XX
 XX 06-APR-1999; 99WO-US007680.
 XX
 XX 07-APR-1998; 98US-0080878P.
 PR 15-MAY-1998; 98US-0085743P.
 XX
 XX (MEDI-) MEDIMUNE INC.
 XX Wizemann TM, Koenig S, Johnson LS;
 PI WPI; 1999-601465/51.
 DR N-PSDB; AAZ31083.
 DR
 XX New pneumococcal proteins useful as vaccines and for diagnosis of
 PT pneumococcal infections.
 PT
 XX Claim 10; Page 72-73; 98pp; English.
 XX
 CC AAY49137-Y49152 are amino acid sequences that are fragments of choline
 CC binding proteins (CBP). The fragments of the protein are the alpha helix
 CC forming parts of the CBPs from Streptococcus pneumoniae. The polypeptides
 CC do not contain the actual choline binding fragment. The polypeptides and
 CC the nucleotide sequences that encode them (AAZ31077-231092) are used in
 CC the invention, which relates to polypeptide truncates of a pneumococcal
 CC surface binding protein containing the highly conserved immunogenic alpha
 CC helical portion and no choline binding portion. The polypeptides are used
 CC as immunogens in a bacterial vaccine. The vaccine can be used for
 CC preventing (immunising) or treating invasive bacterial (especially
 CC pneumococcal) infections, especially otitis media (caused by
 CC S.pneumoniae), sepsis, meningitis and lobar pneumonia infections.
 CC Antibodies raised against the polypeptide are useful for detection,
 CC prevention (passive immunity) and treatment of S. pneumoniae infections.
 CC The vaccines are especially useful in immunocompromised patients, those
 CC with an immature immune system, or patients with an on going pneumococcal
 CC infection. The vaccine avoids unnecessary expense and provides broad
 CC protection against a range of pneumococcal serotypes and it produces an
 CC improved and enhanced effect in preventing bacterial infections
 XX
 SQ Sequence 446 AA;

Query Match 86.8%; Score 2210; DB 2; Length 446;
 Best Local Similarity 100.0%; Pred. No. 1.8e-131;
 Matches 446; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 38 TENEGATQVPTSSNRANESQAEQEQPKLDSERDKARKEVEEVYKIVGESYAKSTKKR 97
 DB 1 TENEGATQVPTSSNRANESQAEQEQPKLDSERDKARKEVEEVYKIVGESYAKSTKKR 60
 QY 98 HTTITVALVNLNINKNEYLKIVESTSESQILMMESRSKVDKAVSKPKDSSSSSSSS 157
 DB 61 HTTITVALVNLNINKNEYLKIVESTSESQILMMESRSKVDKAVSKPKDSSSSSSSS 120
 QY 158 SSTKPEASDTAKPNKPTPEGKVAEAKKKVEAEKAKDOKEEDRRNPTITYKTLELEI 217
 DB 121 SSTKPEASDTAKPNKPTPEGKVAEAKKKVEAEKAKDOKEEDRRNPTITYKTLELEI 180
 QY 218 AESDVEVKKAELELVKVKANPRDEQKIQAEEVESKQAEATRLKKIKTDREAEBAEK 277

Db	181	AESDVEVKKAELVLVKVANS	PRDQKI	KQAEAVESKQAEATRUUKIKTDREEAEEAK	240
Qy	278	RRADAKEQKPKRAKRGV	PGELATPD	KKENDAKSDSVGBETLPSPSLKPEKKVAEAE	337
Db	241	RRADAKEQKPKRAKRGV	PGELATPD	KKENDAKSDSVGBETLPSPSLKPEKKVAEAE	300
Qy	338	KKVBEAKKKAEDQKED	RRNPTNTY	KTLELEIAESDVEVKKAELVLKVEEAKPEPRNEK	397
Db	301	KKVBEAKKKAEDQKED	RRNPTNTY	KTLELEIAESDVEVKKAELVLKVEEAKPEPRNEK	360
Qy	398	VKQAKAEVSKKAETRLUKI	KTDRKKAEEERAKKA	EEDKVKEKPAQOPAPAPAKAEK	457
Db	361	VKQAKAEVSKKAETRLUKI	KTDRKKAEEERAKKA	EEDKVKEKPAQOPAPAPAKAEK	420
Qy	458	PAPAPKPNPAEQPKAEK	PAQQAEE		483
Db	421	PAPAPKPNPAEQPKAEK	PAQQAEE		446

Search completed: September 28, 2005, 18:39:36
Job time : 174 secs

THIS PAGE IS BLANK